Submitted on: 12/08/2014

« ISNI et VIAF transforment le paysage : pour des identités fiables et solides »

French translation of the original paper: "ISNI and VIAF – Transforming ways of trustfully consolidating identities".

Translated by: Pauline Chougnet, Bibliothèque nationale de France, Paris, France.

Le texte de ce document est une traduction française : des differences peuvent exister avec le texte original. Cette traduction n'est fournie qu'à titre indicatif.

Anila Angjeli
Bibliothèque nationale de France, Département Information bibliographique et numérique, Paris, France.
ISNI International Agency
ISNI: 0000 0004 2755 4724

Andrew Mac Ewan
British Library, Department of Collections, London, UK.
ISNI International Agency
ISNI: 0000 0004 2367 7583

Vincent Boulet
Bibliothèque nationale de France, Département Information bibliographique et numérique, Paris, France.
VIAF Council
ISNI: 0000 0003 5723 4556

This is a french translation of "***ISNI and VIAF – Transforming ways of trustfully consolidating identities***" copyright © 2014 by Pauline Chougnet. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License: http://creativecommons.org/licenses/by/3.0/

Résumé :
Cet article présente ISNI et VIAF, deux grandes initiatives d'optimisation du contrôle d'autorité. Elles changent de manière significative le paysage international du contrôle bibliographique en relevant le défi de l'identification fiable des contributeurs dans un réseau global des connaissances en expansion. Les deux systèmes sont conçus sur les principes du contrôle bibliographique universel. De surcroît, ils donnent à ces principes une nouvelle vie en promouvant une économie globale où les données d'autorité jouent un rôle moteur dans la chaîne des liens du dans le web de données (Linked data). L'article montre la complémentarité d'ISNI et de VIAF et encourage les bibliothèques du monde entier à prendre part à ces deux initiatives.

Keywords: VIAF, ISNI, Contrôle d'autorité, identifiant standard international, agrégation de données.

Introduction

Les bibliothèques, et plus largement les institutions patrimoniales, ont traditionnellement concentré leurs efforts sur la collecte et le catalogage des produits de l'activité humaine. La description des personnes liées aux ressources était une activité secondaire, dont l'objectif essentiel était de fournir un accès aux ressources elles-mêmes. Cependant, l'émergence de l'Internet, en tant que phénomène social, nous a finalement sensibilisés à la place centrale qu'occupent les personnes dans nos missions, que ce soit pour leur rôle de producteurs des ressources que nous conservons dans nos collections, mais aussi pour leur rôle d'utilisateurs de ces ressources et potentiellement pour celui de producteurs de connaissances. Nous sommes de plus en plus conscients de notre très forte implication dans une entreprise humaniste vaste, profonde et continue.

Les personnes, les individus comme les collectivités, figurent parmi les principales entités qui permettent d'interconnecter des informations entre des bases de données, des collections et des systèmes et, plus généralement, de les lier sur le web. Dans ce contexte, la longue tradition de création, de maintien et de partage des données d'autorité par les bibliothèques leur est un atout inestimable. Les fichiers d'autorité, où l'on trouve une identification et une désambiguisation fiables, de nombreuses variantes de noms et d'autres informations pour l'identification - justifiées par des sources - dans des normes professionnelles et une structuration fine des données, sont des candidats idéaux pour être intégrés par les nouvelles technologies et pour être réutilisés de différentes façons, y compris en dehors de la communauté internationale des bibliotèques.

Conçu dans les années 1970, le concept de contrôle bibliographique universel (CBU) était révolutionnaire et visionnaire : il s'agissait d'organiser toute l'économie de la production des données bibliographiques et du contrôle qualité à l'échelle internationale par le partage des responsabilités et des bénéfices. Dans le contexte du web de données ouvertes (*Linked Open Data*), alors que leur trafic ne cesse d'augmenter dans le monde, les questions de qualité des données et de confiance sont devenues cruciales et hautement stratégiques. La nécessité de données fiables pour l'identification des personnes devient pressante pour consolider la pertinence du réseau de connections, au moment où les nouvelles technologies développent les interactions et le tissage de liens entre données. Dans ce contexte, les erreurs se propagent, se diffusent, et nuisent à la crédibilité du réseau. Ces erreurs, dès lors qu'elles sont diffusées, sont plus difficiles à repérer et corriger. Les moyens d'appliquer les principes du CBU doivent être redéfinis en tenant compte de ce contexte.

Sont présentés ici VIAF et ISNI, deux initiatives majeures se proposant d'identifier les personnes dans ce réseau global de connaissances en croissance rapide. L'article analyse comment les deux systèmes confortent les principes du CBU. De plus, il démontre que VIAF et ISNI changent de manière significative le paysage international du contrôle bibliographique et qu'ils créent des outils pour les activités de contrôle dans le réseau de connaissances auxquel les bibliothèques participent. Tous deux jouent un rôle important pour l'alignement, l'amélioration et la diffusion des données de bibliothèques, en leur conférant un rôle de pivot dans l'économie globale des connaissances. Les resssources qu'elles créent ont un impact grandissant dans l'économie de la production des données, notamment par le

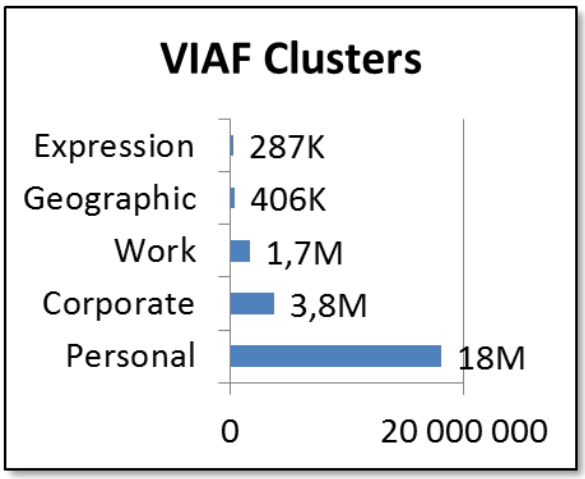
partage du travail – énergivore - de qualité sur les identités. VIAF et ISNI, qui sont complémentaires, contribuent tous les deux à ce paysage émergent.

Contrôle d'autorité et identification

VIAF – le rôle central du contrôle d'autorité

VIAF est une source majeure pour le contrôle d'autorité et devient la référence commune à l'échelle internationale (Bourdon et Boulet, 2013). Centré sur les ressources des grandes bibliothèques de référence – principalement des bibliothèques nationales et des catalogues collectifs à dimension nationale – VIAF rassemble aujourd'hui les données d'autorité de 43 institutions[1]. Outre les données de bibliothèques nationales et de catalogues collectifs, VIAF est aussi adapté à d'autres sources, spécialisées dans des domaines particuliers. VIAF est connecté au domaine encyclopédique : on y trouve les données du Getty et d'autres bases de données spécialisées du domaine universitaire, comme Perseus, hébergé par l'Université Tufts, et le „Syriac reference Portal" de l'Université Vanderbilt, ainsi que des données venant de Wikipédia. VIAF se positionnne ainsi à la croisée des données de bibliothèques et du champ patrimonial au sens large, et a pour intérêt d'accroître la visibilité des données d'autorité dans la longue traîne du web. A plusieurs égards, VIAF est la concrétisation d'un projet depuis longtemps envisagé, programmé et discuté à l'IFLA depuis 1978 (Plassard, 2003).

VIAF couvre un large éventail d'entités identifiées par des autorités. Bien qu'il se soit limité au début aux personnes et aux collectivités, sa couverture n'a pas cessé de s'élargir ces dernières années. Aujourd'hui, il comprend également des oeuvres, des expressions et des noms géographiques. Il sera crucial pour VIAF dans les prochaines années de se concentrer sur ces entités, étant donné que ces activités sont intrinsèquement liées à l'évolution des catalogues et à leur FRBRisation. Le chargement fait par OCLC dans VIAF de données sur des oeuvres et des expressions, basées sur une extraction de données des notices bibliographiques de Worldcat, souligne le rôle que prendra VIAF prochainement dans un domaine si stratégique (Hickey, 2014).



(En février 2014)

Figure 1: Grappes de VIAF– volumétrie selon le type d'entités

[1] En juillet 2014, VIAF inclut 26 grandes bibliothèques, catalogues collectifs et institutions culturelles et de recherche, 9 bibliothèques membres de NACO (Name Authority Cooperative) et 8 autres sources sont en test (bibliothèques, Wikipédia, bases de recherche)

ISNI – LE ROLE DE L'IDENTIFICATION PERENNE

Une autre initiative, ISNI - ISO 27729, répond au besoin déjà ancien, d'un identifiant unique, global, multi-domaines, normalisé et pérenne s'appliquant aux personnes et aux collectivités impliquées dans les contenus créatifs. Il faut remonter aux années 1970 pour retracer l'histoire des initiatives menant à la création d'un tel identifiant (Angjeli, 2012). La mise en oeuvre concrète de l'ISNI, initiative commune à différents acteurs des industries de l'information[2], a été favorisée par les développements récents dans l'économie du numérique. Les identifiants standardisés, pérennes et uniques y sont considérés comme la pierre angulaire pour garantir le plus haut niveau d'automatisation. Rejoignant les préoccupations communes à tous ceux qui gèrent de l'information sur les personnes et les collectivités, comme cela a été développé dans l'introduction, d'autres préoccupations ont nécessité une solution opérationnelle plus ciblée et plus englobante. Plus ciblée, c'est-à-dire qui traite spécifiquement de la question d'un l'identifiant normalisé. Plus englobante, dans le sens où cet identifiant soit universel et multi-domaines, afin d'être utilisé dans des opérations d'échanges internationales entre acteurs de différents domaines.

Dans la communauté des bibliothèques, le groupe de travail sur les spécifications fonctionnelles et la numérotation des notices d'autorités (FRANAR), a été chargé en 1999 d'étudier la question d'un numéro unique pour identifier la même personne, la même collectivité ou n'importe qu'elle entité qui fasse l'objet d'une notice d'autorité[3]. En 2008, ce groupe de travail a publié une étude intitulée „A review of the feasability of an International Standard Authority Data Number (ISADN)" (FRANAR, 2008). Tout en reconnaissant les avantages d'un numéro unique pour éviter les doublons et qui soit utilisable indépendamment de tout language et de tout système, le groupe de travail recommandait à l'IFLA de ne pas poursuivre l'idée de l'ISADN lui-même. A la place, il recommandait de soutenir les développements du groupe de travail sur l'ISNI[4] et le projet VIAF ainsi que de chercher à influencer l'ISNI afin qu'il prenne en compte des intérêts communs à différentes communautés.

Fidèles à cette recommandation, les bibliothèques ont toujours été une force motrice dans le projet ISNI, depuis l'élaboration du standard ISO 27729 à la mise en oeuvre du système ISNI et de son réseau. La CENL (représentée par la British Library et la Bibliothèque nationale de France) et OCLC sont membres de l'Agence Internationale ISNI (ISNI-IA), l'instance de gouvernance de l'ISNI[5]. OCLC gère la base de données centrale[6] et assure la fonction d'agence d'attribution de l'ISNI. Des experts en données d'autorité de la BnF et de la BL sont responsables de la qualité de la base : ils constituent l'équipe qualité ISNI.

[2] Ceux-ci comprennent les bibliothèques, les sociétés de gestion de droits, les acteurs de l'interprofession du livre, les agrégateurs et les fournisseurs de services.

[3] La lettre de mission du groupe FRANAR comprenait deux autres tâches : définir les spécifications fonctionnelles des notices d'autorité (une mission qui a conduit à l'élaboration en 2009 de FRAD - Functional Requirements for Authority Data : a conceptual model) ; opérer comme organe de liaison officiel de l'IFLA avec d'autres groupes intéressés par les fichiers d'autorité http://archive.ifla.org/VII/d4/wg-franar.htm

[4] Le groupe de travail ISO pour l'élaboration du standard ISNI a débuté en 2006. Quand l'étude de FRANAR sur l'ISADN a été publiée, le standard ISNI était en train d'être élaboré.

[5] ISNI-IA est constituée de 6 organisations internationales, de multiples domaines, publiques et privées. Les membres de l'ISNI-IA sont : la CENL, la CISAC (International Confederation of Authors and Composers Societies), l'IFROO (International Federation of Reproduction Rights Organisations), ProQuest et SCAPR (Societies' Council for the Collective Management of Performers' Rights). ISNI-IA mène la stratégie générale de l'ISNI pour le développement de son système et de son réseau. Il prend des decisions en matière de politique des données, d'opérations techniques, économiques et stratégiques.

[6] La base de données ISNI est accessible depuis le site official d'ISNI http://www.isni.org

La base de données ISNI, pour construire un système d'identifiants multi-domaines, intègre des données de différents contributeurs de données, incluant des bibliothèques et d'autres sources. Ils sont aujourd'hui plus de 70 (en juin 2014). Outre les 41 contributeurs de VIAF, 30 autres organisations ont chargé leurs données dans ISNI.

Objectifs

VIAF

VIAF est conçu comme un outil de référence pour les bibliothèques, les archives et les musées du monde entier. Ses objectifs principaux sont le partage de données, la réduction des coûts de catalogage (rendue possible par la réutilisation du travail fait par d'autres), la fourniture de données d'autorité dans des languages et systèmes d'écritures nécessaires aux utilisateurs. Les utilisateurs peuvent réutiliser les données librement, grâce à une licence de données ouvertes ODC-By. Les données sont exposées dans divers formats d'export (RDF, MARCXML, JSON) qui sont employés dans la communauté des bibliothèques aussi bien qu'en dehors. VIAF représente pour ses membres une opportunité inédite de construire une communauté d'intérêts. Chaque membre de VIAF prend part au conseil VIAF. Il constitue un outil de la communauté en donnant une nouvelle visibilité aux données de ses membres à l'échelle internationale. Cela a toute son importance pour pouvoir juger de la qualité des données en fonction de leur provenance : VIAF met en relief et valorise les domaines de compétence de chacun des contributeurs, comme les bibliothèques nationales pour leur propre couverture nationale.

D'autre part, VIAF permet à ses membres d'améliorer leurs propres données avec des informations dans des secteurs où ils déploient moins d'efforts ou qui n'entrent pas dans leur domaine de compétence. Chaque partenaire de VIAF bénéficie de l'expertise des autres partenaires. VIAF joue également un rôle important pour connecter les données à d'autres projets internationaux, comme DBpedia et ISNI.

ISNI

L'ambition première d'ISNI est de fournir un identifiant pour une gestion interopérable des informations sur - et relatives à - des personnes et des collectivités dans l'environnement numérique. Son utilisation est cruciale, non seulement pour faciliter la recherche et la découverte de ressources mais aussi pour fluidifier les transactions. ISNI fonctionne quels que soient la langue et le système utilisés et il est également conçu indépendamment de tout secteur et de tout territoire. En d'autres termes, ISNI est un identifiant multi-domaines, une passerelle entre les données. Le cahier des charges pour la conception de l'identifiant passerelle et multi-domaines qu'est l'ISNI associe en réalité plusieurs objectifs :

- **Un identifiant unique pour tous les secteurs**. Les personnes et les collectivités peuvent s'impliquer dans des activités très diverses et dans différents secteurs, se déplacer d'un lieu vers un autre, créer dans plus d'une langue, être publiées dans d'autres pays que leur pays d'origine ou de résidence, etc. Un seul identifiant ISNI est attribué à chacune des identités publiques (voir ci-dessous) d'une personne ou collectivité, quel que soit le secteur ou le rôle qu'elle a tenu dans ces contenus créatifs.

- **Un identifiant pour tous les usages.** ISNI, identifiant normalisé de l'ISO, est conçu pour répondre aux besoins de toutes les parties prenantes de l'industrie de la documentation, dans le monde entier, indépendamment de leurs secteurs d'activités, y compris les bibliothèques, archives et musées, les éditeurs, les acteurs de la distribution, les sociétés de gestion de droits, les organismes de recherche, les aggrégateurs de données, les centres de traitement de données, et d'autres encore.

- **Un identifiant global, de haut niveau, *versus* des identifiants locaux**. ISNI est un identifiant interopérable dans sa conception : l'essentiel de ses missions est de faire le lien avec d'autres standards ou identifiants propriétaires (...)" (LCC, 2014a). ISNI fonctionne à un haut niveau d'interopérabilité. Il ne remplace pas nécessairement les identifiants utilisés au niveau local, dans des réseaux particuliers ou des consortiums, qu'ils soient ouverts, publics ou propriétaires. De tels identifiants sont nécessaires pour la gestion interne des systèmes d'information : des circuits de données ou des opérations d'échanges ont pu se construire à partir d'eux. En intégrant les identifiants locaux de différents contributeurs de données dans les processus de consolidation pour une identité, ISNI relie ces identifiants à l'identifiant ISNI de haut niveau, ce dernier étant attribué à l'identité publique (et non à la notice), comme le schéma n°2 le représente. La base de données, gérée de manière centralisée, fournit l'infrastructure qui garantit la consolidation, la maintenance, la mise à jour et la diffusion de tels liens. De cette manière, ISNI joue un rôle d'identifiant passerelle.

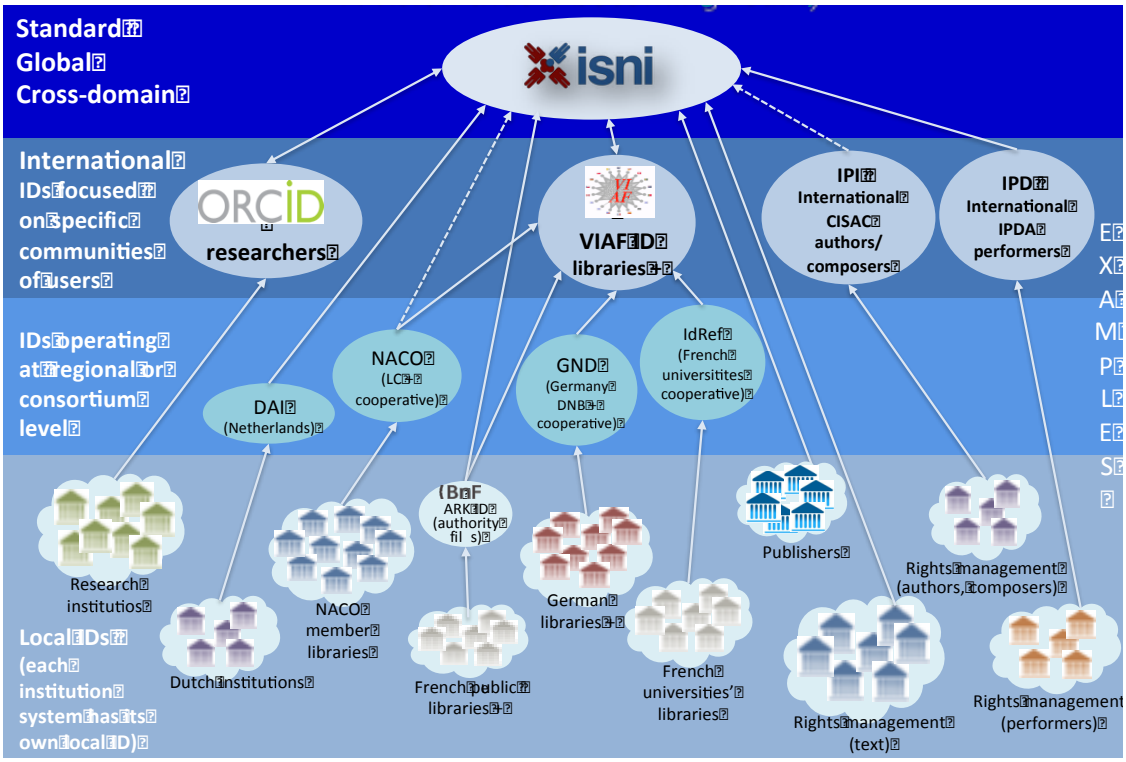


Figure 1: Différents niveaux d'identifiants associés aux personnes et collectivités

ISNI a pour ambition de se positionner comme un identifiant pivot, valable dans tous les domaines, en donnant une place centrale aux fichiers d'autorité, tout en dépassant les limites de leurs frontières traditionnnelles.

Philosophie
VIAF
VIAF permet aux données d'autorité de divers contributeurs d'être mises en relation les unes aux autres. Cette mise en relation donne la possibilité à chacun des partenaires de corriger et de consolider les données et d'établir des associations fiables. VIAF met en relief des points de convergence et de divergence entre les données des membres. Il repère ainsi également des données qui doivent être vérifiées et corrigées. Par exemple dans la grapppe représentée ci-dessous, la date de naissance de Martynas Mažvidas doit être déterminée :

Figure 3: La grappe deVIAF pour Martynas Mažvidas': les données fournies divergent quant à la date de naissance

Tels sont les principes fondamentaux et la philosophie de VIAF. VIAF lui-même ne crée aucune donnée. Il ne fait qu'associer et rassembler des données fournies par ses membres. Il est important de noter que VIAF reste neutre vis-à-vis des politiques de catalogage de ses partenaires et de leurs formats de production. VIAF prend en compte différents modèles de données et règles de catalogage. Il construit des ponts, sans imposer des règles ou des bonnes pratiques à ses partenaires. Ses membres utilisent des standards nationaux ou RDA, utilisent le format Marc 21, Unimarc ou même MADS. VIAF adopte le point de vue de la donnée elle-même, il indique uniquement si la donnée d'un des partenaires est considérée comme la même ou suffisamment proche de celle d'un autre. En cas de doute, il suspend son jugement, et crée une autre grappe. D'une certaine manière, VIAF a une approche pyrrhonienne de la donnée.

Ces principes sous-tendent le mode opératoire de VIAF. Des algorithmes ont été spécialement conçus pour traiter les données fournies par les partenaires, données qui constituent la matière brute de VIAF. Ces algorithmes prennent en compte le travail effectué par les partenaires dans leurs institutions. VIAF utilise exclusivement des traitements automatiques pour rapprocher les données. Il est possible d'intervenir manuellement uniquement de manière exceptionnelle. VIAF ne fournit aucun travail d'expertise sur la qualité des données mais repose sur le travail qualité effectué par ses partenaires. C'est uniquement par comparaison et regroupements de données de qualité que VIAF est capable de créer une vraie valeur ajoutée à la donnée. Si les données entrantes sont suffisantes et de grande qualité, le résultat de l'appariement est fiable. Si les données fournies sont de mauvaise qualité, pauvres ou indifférenciées, le travail de rapprochement est compromis et le risque est élevé de créer de nouvelles grappes en doublon pour la même entité, ce qui est une conséquence indésirable.

Une fois créées, qu'arrive-t-il aux grappes de VIAF lors des mises à jour et l'entrée de nouvelles sources ? Les grappes de VIAF sont-elles stables ? Si aucun nouveau membre n'apporte ses données dans une grappe de VIAF ou si aucune des notices d'autorité composant la grappe n'est mise à jour[7], la grappe reste stable. Mais si les données d'un nouveau membre sont ajoutées à une grappe, ou si une nouvelle version d'une notice déjà présente (corrigée ou complétée avec d'autres informations) est chargée, l'équilibre qui constituait la grappe peut être modifié. Le calcul algorithmique de ce nouvel équilibre peut démontrer que la précédente association était fausse. Cela signifie que toutes les données présentes dans une grappe sont à nouveau appariées, en tenant compte des nouvelles informations. Soit ce nouveau traitement se justifie et mène à une amélioration de la qualité de la grappe, soit il produit un effet de bord indésirable, comme un mouvement injustifié d'une notice d'une grappe à une autre.

C'est pourquoi, la question de l'instabilité des grappes de VIAF ne peut être considérée en bloc. La véritable question est : qu'y a-t-il de stable et pour quoi faire ?

Quant à l'identifiant de VIAF, il est attribué à une grappe de données, y compris quand celle-ci ne contient que les données d'un seul partenaire. L'identifiant de VIAF ne se rapporte pas à l'entité elle-même (personne, collectivité, oeuvre,...) mais à une certaine grappe de données. Cependant, bien qu'il identifie une grappe de données maintenue par le système VIAF, cet identifiant a une valeur internationale, car le domaine de couverture de VIAF est international.

ISNI

Les principes fondamentaux d'ISNI sont largement inspirés, en théorie et en pratique, des principes du contrôle d'autorité des bibliothèques.

Le concept central de l'ISNI est celui de „l'identité publique", ainsi défini par la norme ISO 27729 „les identités utilisées publiquement par les Parties, impliquées dans les industries de contenus : dans la chaîne de création, production, gestion, et distribution". Bien que le terme ne vienne pas du vocabulaire des bibliothèques, le concept lui-même désigne quelque chose se rapprochant de ce qui fait l'objet d'une notice d'autorité pour une personne ou une collectivité.

En réalité, c'est FRAD qui a éclairci la théorie en vigueur dans les bibliothèques en indiquant explicitement que les personnes et les collectivités traitées dans les catalogues de bibliothèque sont des „entités bibliographiques" ou des „identités bibliographiques". „En tant que telles, elles reflètent les constructions intellectuelles ou des concepts constitutives dans la création des catalogues de bibliothèques (...)" (IFLA FRANAR, 2013). Selon les règles appliquées dans un catalogue, une notice d'autorité pour une personne ou une collectivité peut ou non identifier une personne réelle ou une collectivité. „Les auteurs peuvent être considérés, dans certaines circonstances, comme représentant plus d'une identité bibliographique, et dans ce cas, une instance spécifique de l'entité bibliographique personne peut correspondre à une *persona*, adoptée par un individu, davantage qu'à l'individu *per se*." (IFLA FRANAR 2013). Un exemple caractéristique est celui des pseudonymes. Il faut noter

[7] VIAF reçoit des mises à jour régulières de ses partenaires comprenant des nouvelles créations, des suppressions et des modifications (y compris des modifications faites dans la notice elle-même et des scissions de notices).

également que RDA recommande l'établissement d'identités publiques différentes pour chaque pseudonyme de la même personne[8].

Le standard ISNI adopte toutefois un point de vue strict et établit une distinction claire entre la personne réelle ou la collectivité et son identité publique ou ses identités. Un seul identifiant ISNI distinct est attribué à une seule identité publique. Les métadonnées associées, dont la majeure partie vient de VIAF, aident à la désambiguisation de l'identité publique au sein du système central d'attribution (formes retenues du nom, variantes de nom, dates, pays, éditeur, titres d'oeuvres associées, affiliations, etc.). Des liens sont faits entre les identités publiques d'une personne réelle si, et seulement si, la relation entre ces identités est exprimée publiquement. Dans le cas contraire, pour des raisons de confidentialité, aucun lien n'est fait.

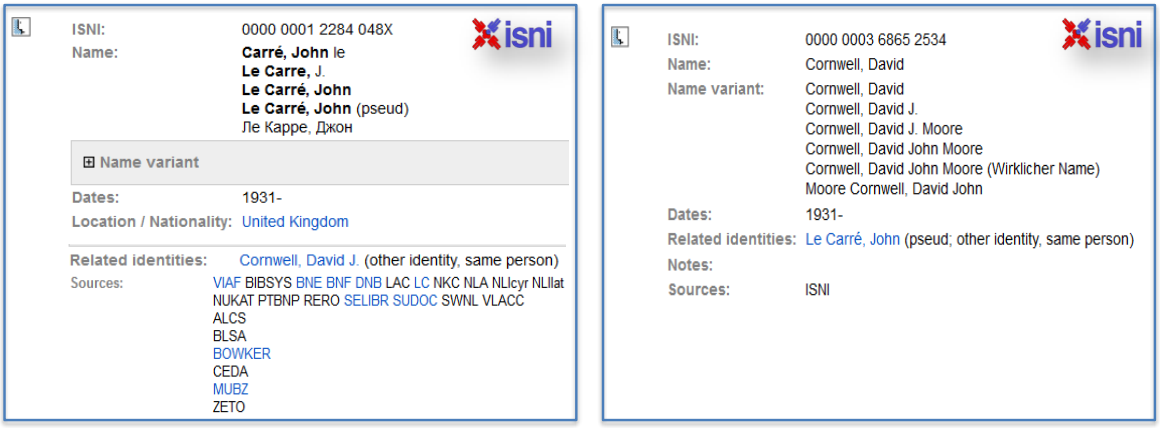


Figure 2: John Le Carré and David Cornwell sont deux identités publiques d'une meme personne. Un identifiant ISNI distinct est attribué pour chacune d'elles.

Le système ISNI est créé directement grâce à l'héritage des bibliothèques. Les données des bibliothèques sont au coeur du système, les données de VIAF ayant été utilisées comme fichier de base lorsque la base de données ISNI a été conçue en 2011. Ces données ont été confrontées à celles de contributeurs extérieurs à VIAF[9]. ISNI traite régulièrement des mises à jour des données de VIAF. Par conséquent, une grande proportion des identités venant des fichiers d'autorités des contributeurs de VIAF reçoit des identifiants ISNI.

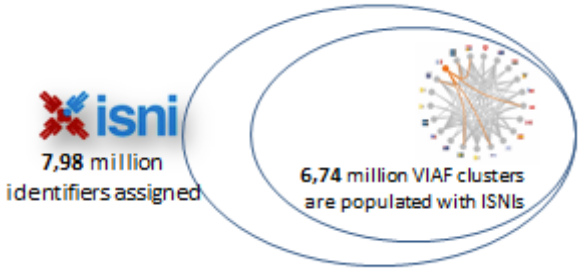


Figure 3: Environ 85% des ISNI attribués contiennent ISNI parmi leurs sources (juillet 2014)

[8] RDA 9.2.2.8 (« Individus ayant plus d'une identité »)

[9] Cela comprend des données sur des livres imprimés, des thèses, des articles, des tables des matières ; des données de sociétés de gestion de droits dans le domaine du texte, de la musique, du théatre ; des données d'organismes de recherche et de sociétés professionnelles et d'encyclopédies en ligne. Pour une liste actualisée des contributeurs de données dans ISNI, voir : http://www.isni.org/content/data-contributors (Consulté le 13 juillet 2014)

Différences entre VIAF et ISNI

Bien qu'ISNI et VIAF aient beaucoup en commun, il existe des différences dans la philosophie et dans la mise en oeuvre de chaque système, qui viennent des objectifs fondamentaux de chacun d'eux. Nous soulignerons brièvement ci-dessous les critères principaux qui constituent la différence fondamentale entre VIAF et ISNI (Mac Ewan, Gatenby, Angjeli ; ISNI-IA 2013) :

- D'une part, ISNI est doté de son propre organe de gouvernance et d'une infrastructure permanente de contrôle, qui est indépendante des pratiques particulières ou de l'arrière-plan culturel de chacun de ses contributeurs de données. Le modèle des données, et le plan de développement du système de l'ISNI sont discutés au sein de l'ISNI-IA. Le contrôle réside dans les méthodes automatiques et dans l'expertise humaine. Les algorithmes sont développés et améliorés par l'équipe ISNI d'OCLC à Leyde et des experts en données d'autorité de la BnF et de la BL, constituant l'équipe qualité, effectuent des tests sur la base. L'équipe ISNI d'OCLC et l'équipe qualité BnF-BL travaillent ensemble au quotidien.

  Pour ce qui est de VIAF, il reste proche de ses partenaires. Il se donne pour objectif d'être ouvert à divers standards, pratiques ou traditions de catalogage et d'être un pivot pour les fichiers d'autorité de ses partenaires. Par cette approche, VIAF souligne les points communs dans les données provenant des partenaires sans exclure aucune méthode de modélisation ou de normalisation. Le conseil VIAF, avec des représentants de chaque contributeur, donne des recommandations à OCLC pour l'amélioration de VIAF ; il représente une communauté de discussion sur les données d'autorité à l'échelle internationale.

- Tandis que VIAF reste neutre vis-àvis des différences de catalogage de ses contributeurs de données[10], ce n'est pas le cas pour ISNI. Il a une politique stricte qui découle de son objectif premier : consolider chaque identité publique et attribuer à chacune un identifiant ISNI unique et fiable qui puisse dès lors être diffusé largement et cité en toute confiance. Bien qu'ISNI ait commencé en utilisant les mêmes algorithmes que VIAF pour les rapprochements et les fusions, ces algorithmes ont rapidement été modifiés et ont été précisés et renforcés régulièrement ces quatre dernières années pour mieux correspondre aux objectifs de l'ISNI[11].

Pour prendre une image, VIAF peut être vu comme une cour de récréation pour des données hétéroclites et diverses dont ISNI serait le surveillant. Il n'y a pas de cour de récréation sans surveillant et pas de surveillant sans cour de récréation.

Pour ces raisons, la relation entretenue avec les sources de données est différente chez ISNI et chez VIAF. ISNI gère les notices des identités publiques et met à disposition des outils de production, utilisés par l'équipe qualité et par les membres de l'ISNI, qui utilisent une interface dédiée. Chaque membre de l'ISNI peut s'assurer que ses données sont bien alignées

[10] Par exemple, certaines bibliothèques créent des notices d'autorité séparées pour les différentes identités publiques d'une même personne ; d'autres créent une unique notice d'autorité qui regroupe toutes les identités de la même personne réelle, en prenant un des noms comme forme retenue du nom et les autres noms des autres identités comme des variantes.

[11] ISNI procède à des scissions et à des fusions de données venant de VIAF quand cela est nécessaire et peut protéger les données qui ont été corrigées manuellement.

avec celles des autres. Cela implique que la base de données soit autonome vis-à-vis de ses sources de données lorsqu'une décision doit être prise pour l'établissement des identités publiques. Afin d'assurer l'identification, ISNI peut décider de scinder ou fusionner des données d'un même contributeur et un identifiant ISNI n'est attribué que si les données fournies et le rapprochement fait sont suffisamment fiables (pour en savoir plus, voir le paragraphe sur le CBU). „En substance, ISNI prend à VIAF le concept d'un fichier international d'autorité, mais il fournit une boîte à outils pour le retravailler, donne accès à un plus large éventail de sources et donne la possibilité de faire des chargements réguliers en masse de données d'autorités" (Mac Ewan, 2013). En d'autres termes, la base de données ISNI a le potentiel d'évoluer vers un fichier d'autorité virtuel partagé, comparable à d'autres fichiers d'autorités en réseau comme LC/NACO mais avec des objectifs et une couverture plus universels.

CBU, VIAF et ISNI : du partage au contrôle, du contrôle au partage

Dans le nouvel environnement des données, où les bibliothèques font face aux problématiques du web sémantique et au paradigme des données ouvertes, la visée noble du contrôle bibliographique universel n'est pas un paradis perdu ; au contraire, elle doit seulement être réinvestie et développée pour s'adapter à ce nouveau contexte.

La nouvelle déclaration professionnelle de l'IFLA résume la problématique de cette manière : „Dans les années 1990, on admettait qu'il était important d'avoir des données dans une langue et dans un système d'écriture accessibles aux utilisateurs, mais aussi qu'il fallait respecter la diversité culturelle chez les utilisateurs du monde entier. Ce ne sont pas nécessairement des mesures contradictoires, comme le Fichier international d'autorités virtuel (VIAF) le démontre. Les technologies du web offrent de nouvelles possibilités de partage des données à une échelle internationale et au delà du monde des bibliothèques, mais elles révèlent également un besoin de données d'autorités et de confiance. Ainsi, le concept de CBU est-il toujours valable et mérite que l'IFLA continue à le porter et le promouvoir." (IFLA, 2012)

On voit dans les exemples de VIAF et d'ISNI que leurs réalisations et leurs projets reposent sur les principes du CBU qu'elles promeuvent en les mettant en pratique d'une manière innovante, dans un environnement technologique mouvant.

VIAF, les bibliothèques et le CBU

Le principe fondamental du CBU est le partage. Chaque acteur national, chaque agence bibliographique nationale a une responsabilité particulière pour la production des données d'autorité relative à son domaine de compétence nationale et les met à disposition pour que d'autres acteurs à l'échelle internationale, y compris en dehors du monde des bibliothèques, les réutilisent. La déclaration de l'IFLA sur le CBU précise plus loin :

> „Les ABN (Agences bibliographiques nationales), dans le cadre de leurs activités de création de données bibliographiques de référence, ont la responsabilité de documenter les accès autorisés pour les personnes, familles, collectivités, noms géographiques et titres uniformes d'oeuvres concernant leur pays et de mettre ces données à disposition des autres ABN, bibliothèques et autres communautés (par exemple les archives et les musées)" (IFLA 2012).

VIAF est l'héritier direct du principe de partage du CBU. Chaque ABN fournit ses données, reconnaissables facilement à l'affichage public des grappes par leur drapeau national. VIAF juxtapose différents domaines nationaux et, peut être considéré comme un pivot pour le CBU sur les données d'autorités, où chaque ABN peut assumer son rôle de source de référence. A chaque nouveau partenaire, de nouvelles opportunités sont offertes pour mettre en valeur des aires spécifiques de compétence et de responsabilité, non seulement pour les ABN, mais aussi pour d'autres contributeurs de données, y compris des institutions de recherche, des musées, des archives ou autres.

Les partenaires de VIAF ne lui fournissent pas seulement des données en lien avec leur domaine national mais leurs jeux de données dans leur ensemble. En général, pour donner accès à ses collections, chaque institution étend le champ de son fichier d'autorités au-delà de sa propre aire de compétence. Cela permet à VIAF de ne pas être seulement une juxtaposition de différents jeux de données spécialisés des ABN, tous relatifs à leur propre domaine de spécialité, mais d'être aussi le terrain d'une large confrontation et comparaison, pour une complémentarité des données sur une base démocratique. Les principes du CBU sont donc appliqués d'un point de vue large, les domaines de compétence reconnus ne sont pas seulement juxtaposés de manière statique mais sont comparés de manière dynamique, et la diversité culturelle est promue. VIAF a pour avantage de révéler l'originalité des données fournies par chacun des partenaires. En ce sens, tandis que les données fournies par le partenaire reconnu comme la „source d'autorité" dans un domaine national spécifique ou un autre domaine, sont les plus susceptibles de constituer le coeur d'une grappe, d'autres données, fournies par d'autres partenaires (une autre bibliothèque nationale, un catalogue collectif d'un pays ou une petite institution spécialisée) enrichissent la même grappe de VIAF. C'est ce que l'on peut observer dans l'exemple d'Honoré de Balzac :

Figure 6: La grappe d'"Honoré de Balzac" dans VIAF

L'ABN française (la bibliothèque nationale de France et le SUDOC) fournit le coeur des données pour l'auteur français, ainsi que la forme officielle et nationale du nom, la bonne orthographe et la structure de la forme retenue, selon les règles nationales, comme le recommande le document de l'IFLA „Names of persons : national usages for entry in catalogues[12]". Les autres partenaires utilisateurs de VIAF sont encouragés à utiliser cette forme nationale du nom. Cependant, chaque bibliothèque ou fournisseur de données a son propre environnement de catalogage et a des compétences dans des aires qui peuvent compléter le savoir collectif. Par exemple, les bibliothèques qui cataloguent dans un système d'écriture particulier, associé à leur environnement linguistique, comme la forme hébraïque de la Bibliothèque nationale d'Israël, ou la forme arabe de la bibliothèque nationale du Liban et de la Bibliothèque d'Alexandrie, peuvent compléter la grappe avec ces formes du nom

[12] http://www.ifla.org/files/assets/cataloguing/pubs/names-of-persons_1996.pdf

dans différents systèmes d'écriture. Il est logique que la Bibliothèque nationale de France ne fournisse pas de telles formes du nom pour ses identités nationales. D'autant plus que, pour les langues et écritures hébraïques et arabes, les formes transmises par ces bibliothèques, seront celles qui feront autorité comme elles sont dans leur domaine de compétence. D'autres partenaires peuvent également contribuer en donnant des éléments supplémentaires à la grappe, qui correspondent à leurs propres règles, comme un qualificatif supplémentaire („french writer" pour le Getty). De cette manière, les données fournies par les ABN, considérées comme la source faisant autorité pour une identité donnée, sont consolidées et enrichies par les autres partenaires de VIAF. VIAF fournit ainsi un modèle où les données centrales d'une ABN, considérée comme la source faisant autorité pour un type donné d'identité, devient le noeud central autour duquel gravitent les données des autres partenaires.

Une autre caractéristique de VIAF est qu'il révèle des différences entre les fournisseurs de données. Soit les différences sont considérées comme mineures et n'empêchent pas les données de se rapprocher, soit elles sont considérées comme significatives et empêchent la fusion dans une même grappe. Par exemple, le juriste allemand Moritz August von Bethmann-Hollweg (1795-1877) entrant dans les attributions de la bibliothèque nationale d'Allemagne, c'est elle qui fournit la forme nationale et officielle du nom, ainsi que les dates biographiques.

Figure 7 : La grappe de VIAF pour Moritz August von Bethmann-Hollweg (1795-1877) avec la forme du nom de la DNB au dessus

Deux bibliothèques, la Bibliothèque nationale de France et le SUDOC n'ont pas été intégrées dans la grappe de Moritz August von Bethmann-Hollweg montrée ci-dessus à cause d'une erreur sur la date de naissance[13]. Celle-ci a eu pour conséquence la création d'une autre grappe de VIAF, distincte constituant un doublon pour cette même identité (voir la figure 8).

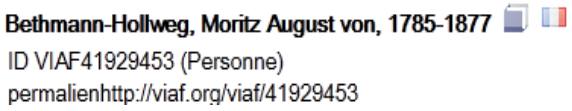


Figure 4 : Une autre grappe de VIAF, un doubloon pour la meme identité, due à une divergence dans la date de naissance

Cet exemple montre comment VIAF permet aux données des autres contributeurs d'être comparées avec les données fournies par l'institution de référence du point de vue du CBU. Dans le cas ci-dessus, la Bibliothèque nationale d'Allemagne étant considérée comme la source de référence pour les personnes, collectivités ou oeuvres allemandes.

D'autre part, VIAF est un fichier vivant et dynamique : il intègre périodiquement des mises à jour et de nouvelles données créées par les institutions. De ce fait, il est un outil

[13] Le SUDOC (Système universitaire de documentation) utilise les autorités de la BnF comme fichier source pour la création de ses propres données d'autorité.

d'interopérabilité et de dialogue entre tous les membres de VIAF, leur servant à améliorer constamment leurs données. Cet exemple montre également que les principes du CBU ne consistent pas uniquement à fournir une forme du nom et des données de référence. La dynamique de VIAF est une incitation au dialogue entre les institutions au sein de la communauté des bibliothèques et au-delà. Les bibliothèques et les institutions patrimoniales sont fortement encouragées à rejoindre VIAF pour partager leurs données avec d'autres contributeurs, mettre en valeur leurs atouts dans un environnement international et tirer davantage parti du travail fourni par les autres.

ISNI – étendre les principes du CBU

Nous nous intéresserons ici à la façon dont ISNI, qui se fonde sur les principes du CBU décrit ci-dessus pour VIAF, approfondit les notions d'"Universel" et de „contrôle", qui lui sont inhérents.

Comme nous l'avons décrit plus haut, ISNI avait à remplir des objectifs différents de ceux de VIAF. Il avait à fournir des solutions pérennes pour que l'identifiant puisse être attribué de manière fiable et pertinente aux identités publiques.

ISNI a créé des règles et des procédures pour assurer un niveau de qualité et de précision nécessaire à la gestion des identités, qui peuvent être résumées brièvement de la manière suivante :

- Un ISNI n'est attribué que lorsqu'un seuil de confiance est atteint : 1) les données elles-mêmes sont suffisamment bonnes pour l'attribution (les données pauvres ou indifférenciées, par exemple, ne sont pas autorisées) et 2) le résultat de l'appariement est élevé. Une fois attribué, le regroupement des sources reste stable et les autres données qui correspondent à ce regroupement doivent, soit rejoindre le regroupement, soit être signalées comme un problème potentiel si elles repèrent plusieurs notices avec lesquelles s'apparier. (Mac Ewan, 2013)

- Les sources de données dans ISNI sont très hétérogènes : les chargements de données de bibliothèques côtoient d'autres données. Un ISNI est attribué uniquement si trois sources de données convergent 1) Une grappe de VIAF est éligible pour attribution uniquement si elle rassemble les données venant d'au moins trois sources 2) une grappe de VIAF comprenant moins de trois sources n'aura un ISNI que si elle s'apparie avec une ou plusieurs sources autres que VIAF ;

- En plus de ces critères, ISNI a développé une règle et des mécanismes pour une attribution à un „Nom unique" („Unique name assignment") et à une „Source unique" („Single source assignment"). La procédure du „Nom unique" assure l'attribution des ISNI à des identités qui sont uniques dans toute la base de données, c'est-à-dire qui ne s'apparient avec aucune autre source mais sont jugées suffisamment fiables. La procédure de la „Source unique" est employée avec beaucoup de précautions pour permettre à des fichiers provenant de sources de référence d'obtenir un statut de „fichier de base" dans le système ISNI, au motif que son émetteur est un vecteur stratégique pour la diffusion de l'identifiant ISNI. Dans ces cas, un effort de développement est déployé pour éviter les cas de doublons qui pourraient en résulter.

- ISNI a également étendu ses techniques d'alignement afin de permettre des regroupements progressifs, en se fondant sur l'appariement croisé des identités figurant dans les notices bibliographiques. Cela permet l'attribution d'ISNI aux auteurs des articles de revues, et d'autres contributeurs de données, qui se trouvent en dehors du cadre des fichiers d'autorités tradtionnels, mais qui sont présents (featuring) dans d'autres bases de données, provenant ou non des bibliothèques14. L'appariement reste fondé sur la concordance d'au moins deux sources distinctes.

Les schémas 7 et 8 ci-dessus montrent que les deux grappes de VIAF concernant le juriste allemand Moritz August von Bethmann-Hollweg (1795-1877) ne devaient pas être fusionnées dans la base ISNI, du fait de leur divergence pour la date de naissance, bien qu'il s'agisse visiblement de la même identité dans les deux. Ils démontrent également que, alors que la grappe représentée dans le schéma 8, contenant une erreur dans la date de naissance, continue d'exister dans les données sources (et de ce fait dans VIAF), elle a été rejetée par les critères d'appariement d'ISNI, qui a jugé que les données n'étaient pas suffisamment fiables. Par conséquent, aucun ISNI n'a été attribué à la deuxième grappe. L'unicité de l'identifiant pour la même identité est dès lors assurée.

Actuellement (juin 2014), 7, 98 millions d'identifiants ISNI, correspondant à ces critères de qualité, ont été attribués. La base de données contient en fait plus de 9, 6 millions d'autres notices qui ne remplissent pas encore les critères d'appariement.

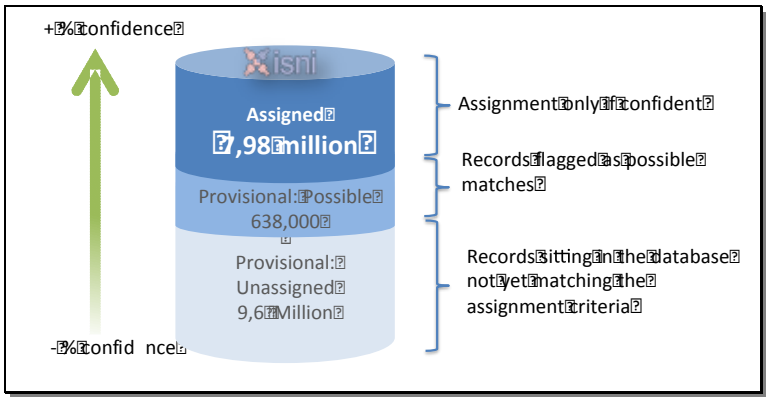


Figure 5 : La base de données ISNI – chiffres – Juin 2014

ISNI doit fournir des solutions pérennes pour que la maintenance et la diffusion de l'identifiant au sein du réseau des contributeurs et partout sur le web soient fiables et pertinentes. Pour cela, un système de notifications automatiques est mis en place entre l'ISNI et ses contributeurs de données. Des notifications mensuelles sur l'attribution, la fusion, la scission ou les suppressions sont envoyées aux contributeurs. De cette manière, le travail effectué par l'équipe qualité d'ISNI rend directement service aux contributeurs. L'une des leçons apprises durant ces quatre dernières années au sein du système ISNI est qu'il est impossible de concevoir une économie globale de partage sans compter sur une participation active des contributeurs au travail qualité. L'objectif des notifications n'est ainsi pas

[14] Plus de 70 000 ISNI ont été attribués dans la base ISNI aux auteurs de thèses de la base EThOS (Electronic Theses Online Service) http://ethos.bl.uk/ . 230 000 ISNI ont été attribués aux auteurs d'articles de revues de la table des matières électronique de la BL «(BL's Electronic Table of Contents http://zetoc.mimas.ac.uk/

seulement de diffuser les ISNI et d'informer les contributeurs sur des changements éventuels concernant leurs données, mais également de solliciter des corrections de données à la source, au sein de l'institution où elles ont été produites. Les agences d'enregistrement et les membres d'ISNI[15], au titre de contributeurs de données privilégiés, ont également accès à une interface de catalogage qui leur permet de participer activement à l'enrichissement et à la correction de notices afin d'assurer un bon résultat d'attribution. Toute l'économie de la production et de la diffusion des données en est transformée, avec un „Contrôle" dans les deux sens. Pour les bibliothèques, s'impliquer dans ISNI (directement ou par le biais de VIAF) revient à placer leurs efforts traditionnels de contrôle d'autorité dans un nouveau contexte, à leur conférer un poids plus stratégique et à s'ouvrir à de nouvelles possibilités de partenariats.

Le travail avec une grande diversité de fournisseurs de données prouve au quotidien la pertinence des grands principes du CBU, selon lesquels le travail de consolidation des identités doit être une responsabilité partagée, en fonction des domaines de compétence de chacun des fournisseurs de données (en confiant par exemple aux ABN la qualité des identités relevant de leur domaine de compétence, aux contributeurs spécialisés la qualité des identités dans leur domaine de spécialité, etc). ISNI, depuis ses débuts, a été l'un des précurseurs de cette approche, puisqu'il attribue aux fournisseurs de données un taux de confiance, en fonction de la provenance de leurs données, et en fonction de leur probable proximité avec le sujet identifié. Le degré de confiance le plus élevé est attribué aux fournisseurs de données qui reçoivent leurs informations directement de la personne ou de la collectivité représentée par cette identité. Quand, pour une certaine identité, des données sont fournies par les bibliothèques, trois niveaux de confiance sont définis, du plus bas au plus élevé :

1. La notice d'autorité est validée par une ou deux bibliothèques
2. La notice d'autorité est validée par trois bibliothèques ou plus
3. La notice d'autorité est validée par trois bibliothèques ou plus ET le pays de provenance d'une de ces bibliothèques correspond à la nationalité de l'identité.

Cela démontre encore les efforts déployés par ISNI pour traduire et mettre en oeuvre les principes du CBU. Mais pour que cela soit complètement fonctionnel, le travail devrait être organisé avec une participation active des bibliothèques elles-mêmes, par une répartition des responsabilités et une définition des circuits.

Il faut noter que l'infrastructure de contrôle de l'ISNI décrite dans cet article est complétée par un mécanisme de correspondance avec le public. L'équipe qualité gère les messages envoyés par le public, en appliquant le principe du „production participative contrôlée" („controlled crowsourcing"). Les suggestions sont analysées et appliquées par l'équipe qualité et des notifications sont envoyées aux utilisateurs pour les informer du résultat de leur demande. De manière caractéristique, ce procédé peut attirer soit les individus qui font eux-mêmes l'objet de la notice dans la base ISNI ou bien des spécialistes dans des domaines particuliers qui veulent améliorer la qualité de la base de données.

[15] Pour plus d'informations sur les agences d'enregistrement et les membres d'ISNI, voir http://www.isni.org/content/isni-community

Les représentants de tous les domaines gérés par la base ISNI (les membre du bureau des directeurs, et les autres contributeurs) reconnaissent de manière unanime l'expertise des bibliothèques dans la gestion des identités. La réputation de l'équipe qualité d'ISNI est déjà bien établie, mais il est nécessaire de renforcer l'équipe. Ici encore, les principes du CBU ouvrent de nouvelles perspectives. Du partage au contrôle, du contrôle au partage, les autres bibliothèques sont invitées à participer au travail de qualité avec l'équipe ISNI et à y contribuer, en fonction de leur domaine d'expertise, afin qu'elles continuent de jouer un rôle moteur dans la gestion des identités, à l'échelle internationale.

L'interopération entre VIAF et ISNI

L'interopérabiblité entre ISNI et VIAF est stratégique, afin de placer les données de bibliothèques au coeur d'un réseau solide et sans précédent d'informations sur les identités, et pour les connecter à des jeux de données extérieurs au monde des bibliothèques.

Bien que les deux systèmes aient des objectifs différents, comme cela a déjà été décrit, ils sont fortement liés l'un à l'autre depuis le début. Mais cette relation doit être encore améliorée pour les rendre pleinement interopérables – ce qui est une condition nécessaire pour une diffusion fiable des ISNI dans VIAF. En fait, la diffusion des ISNI dans les fichiers d'autorité des bibliothèques est déjà implicite, du fait de l'attribution des ISNI aux grappes de VIAF. ISNI est affiché en haut des grappes de VIAF, sous forme de lien cliquable et il est diffusé librement dans les téléchargements des données de VIAF en RDF. Inversement, la diffusion des ISNI, des liens et des URL présents dans la base de données ISNI est la prochaine étape pour mettre en oeuvre les objectifs d'ISNI comme une passerelle entre les données et comme une brique fonctionnelle du Web de données (Linked Data).

Aujourd'hui, des circuits réguliers sont déjà en place entre ISNI et VIAF. ISNI charge les mises à jour de VIAF tous les deux mois et envoie les identifiants ISNI à VIAF une fois par mois. De plus, ISNI améliore constamment les algorithmes d'appariement pour assurer l'unicité de l'identifiant et traite les données par du contrôle qualité manuel (travail effectué par l'équipe qualité). Il détecte des erreurs dans les grappes de VIAF et envoie des notifications régulières à VIAF. Deux grands types d'erreurs sont signalées :

- Des erreurs d'appariements faites par VIAF (y compris des mouvements de grappes, comme cela a été expliqué plus haut)
- Des erreurs dans les données sources des partenaires de VIAF

Si les erreurs d'appariements peuvent être traitées directement par VIAF, les erreurs dans les données sources des partenaires soulèvent la question de l'efficacité des circuits de notifications. A ce niveau, l'implication des partenaires de VIAF eux-mêmes dans le traitement est nécessaire.

Toutes les questions d'interopérabilité VIAF-ISNI sont régulièrement discutées par les structures administratives et techniques des deux systèmes. Au sein du conseil VIAF, un groupe de travail pour l'interopérabilité VIAF-ISNI a été créé. Les questions qui y sont traitées comprennent : l'amélioration des circuits de données entre les deux systèmes, l'intégration des notifications venant d'ISNI dans VIAF, la participation des bibliothèques aux corrections de données, la mise en valeur des données de qualité. Une avancée remarquable de ces efforts de coopération est l'accord selon lequel des notifications d'erreurs doivent être envoyées directement aux partenaires de VIAF, dès lors qu'elles sont détectées dans les données sources et qu'elles sont relevées par l'équipe qualité d'ISNI. Ce circuit vient d'être mis en place et les premiers résultats sont encourageants. Les bibliothèques sont

incitées à devenir des acteurs à part entière dans ce système, en corrigeant leurs données et en collaborant avec l'équipe qualité pour le bénéfice de toute la communauté des bibliothèques et au-delà.

Le résumé ci-dessous montre la complémentarité de VIAF et d'ISNI en ce qui concerne leur domaine de compétence, leur rôle et les fonctionnalités de leurs systèmes respectifs.

Domaine de compétence de VIAF

- Personnes
- Collectivités
- Œuvres / titres uniformes
- Expressions
- Congrès
- Noms géographiques

Nature des données de VIAF

- Toutes les données sont publiques

Rôle de VIAF

- Crée des grappes
- Intègre des notices d'autorité provenant des plus grandes bibliothèques du monde (bibliothèques nationales et bibliothèques de recherche) et d'autres institutions patrimoniales
- Expose et diffuse les données

Le système de VIAF

- Moissonneur
- Mécanisme d'appariement
- Site Internet (interface en 5 langues)
- Système de notifications à envoyer aux fournisseurs de données (en instruction)
- Affichage et téléchargement dans différents formats
- Web de données (Linked Data) et SRU

(Données ouvertes (ODC-By licence)

Domaine de compétence d'ISNI

- Personnes
- Collectivités

(A l'exception des notices pauvres)
(A l'exception des notices indifférenciées)

Nature des données d'ISNI

- Inclut des données privées

Rôle d'ISNI

- Crée un identifiant normalisé ISO
- Intègre des données depuis VIAF et crée des liens avec des sources extérieures à VIAF – une plateforme de liens multi-domaines
- Travaille avec des données hors bibliothèques et fournit des résultats d'appariement (par exemple des données de périodiques, de thèses, etc.)
- Expose et diffuse les ISNI attribués uniquement (pour lesquels on peut être sûr que les notices ont été différenciées et dédoublonnées)

Le système d'ISNI

- Chargements en masse
- Mécanisme d'appariement
  Maintenance des grappes
- Requêtes en ligne par webservices (API)
- Site Internet (en anglais uniquement)
- Services en ligne
  - Messagerie pour les utilisateurs
  - Outil d'édition en ligne pour les membres et les agences d'enregistrement
- Une équipe qualité qui contrôle la qualité de la base de données et corrige les notices
- Notifications envoyées aux fournisseurs de données
- Affichage des données dans différents formats (XML, RDF par une page de négociation de contenus)
- SRU, Web de données (Linked Data) licence de données en cours d'adoption

Conclusion : les principes du CBU dans une nouvelle ère

Comme cela a été démontré tout au long de cet article, la vision soustendue derrière VIAF et ISNI est inspirée par la philosophie du CBU, qui s'inspire elle-même de l'idéal humaniste d'un partage des connaissances et de la promotion de la diversité culturelle tout en permettant une efficacité dans le travail des bibliothèques au quotidien.

Pendant plusieurs décennies, l'état des technologies a limité la réalisation de toutes les potentialités des principes du CBU. VIAF et ISNI, en tirant parti des nouvelles technologies, ont réaffirmé ces principes dans des applications innovantes permettant de nouvelles approches de la question, des applications qui étaient impensables dans un passé pas très lointain. VIAF, qui fédère les fichiers d'autorité des bibliothèques du monde entier, est reconnu comme une brillante réussite. ISNI consolide encore plus les identités, attribue un identifiant ISO, multi-domaine, aux identités publiques des personnes et collectivités et devient par là-même un outil puissant pour construire une chaîne fiable de données liées. Par leurs évolutions techniques et organisationnelles, les deux systèmes qui coopèrent jouent un rôle moteur dans toute l'économie de la production, du contrôle, de la difussion et de la réutilisation des données. A ce titre, VIAF et ISNI renouvellent les principes du CBU et ouvrent une nouvelle ère pour une consolidation collaborative et sans précédent des identités à l'échelle internationale.

De plus, l'expérience du travail dans un environnement ouvert et global pour le partage des données et des connaissances, étend sans nul doute le domaine des principes du CBU, vu que les nouvelles technologies présentent des opportunités que leurs créateurs mêmes n'avaient pas pu prévoir. Mais dans le même temps, ces opportunités s'accompagnent de nouveaux défis.

Le premier défi de taille est la qualité des données sources. Il est évident que plus les données sources sont de bonne qualité, plus les bénéfices apportés par leur rapprochement et leur connection par VIAF et ISNI seront grands. Cela soulève la question de la façon d'encourager les bibliothèques à consacrer des ressources pour le travail qualité sur les données qu'elles partagent en réponse aux bénéfices qu'elles reçoivent grâce à ces nouveaux outils. Une réponse immédiate à cela est que les bibliothèques devraient participer aux deux systèmes pour placer leur production dans un environnement large et assumer pleinement leurs responsabilités dans une économie globale.

Concernant les responsabilités pour la production des autorités, plusieurs questions se posent. L'autorité des ABN, doit-elle être considérée comme définitivement établie, non remise en question ? Celles-ci doivent-elles être les seules insitutions de référence dans ce domaine ? Ou bien doit-on reconnaître la valeur d'autres institutions spécialisées, chacune dans leur propre domaine de compétence ? En rassemblant des données d'une grande diversité de contributeurs, VIAF et ISNI démontrent l'émergence de nouveaux partenaires de confiance pour les autorités, des partenaires pour lesquels la qualité des données est essentielle.

Une autre question se pose : comment doit-on organiser le travail sur les autorités et leur contrôle au niveau local ? La responsabilité des autorités doit-elle être concentrée uniquement sur les ABN ? Des bibliothèques et des institutions territoriales ont des connaissances spécialisées sur les identités relatives à leurs collections spécialisés et leurs fonds locaux ;

d'autres bibliothèques ont des experts dont les connaissances sont pointues sur les identités de leur domaine. Ce savoir spécialisé peut bénéficier à toute l'économie de la production et du contrôle des données d'autorité et permettre de construire un ensemble de données de grande qualité qu'aucun d'entre nous n'aurait pu construire seul. Le rôle des ABN serait alors précisément de mobiliser des compétences, de développer des partenariats, de partager le travail au niveau national pour contribuer au réseau global de partage.

Remerciements
Les auteurs remercient Janifer Gatenby (responsable de la base ISNI à OCLC (Leyde)) pour l'aide précieuse qu'elle leur a apportée.

Bibliographie

Angjeli A. (2012) 'ISNI : un identifiant passerelle', *Documentation et Bibliotheques*, Vol 58 No 3, pp 101-108 [Digitized copy available online from the BnF Open Archives Portal 'Hyper Articles en Ligne'] http://hal-bnf.archives-ouvertes.fr/hal-00803772 (Consulté le 5 juin 2014)

Boulet, V. (2012) 'L'infrastructure de l'information et les besoins des utilisateurs : tout le pouvoir aux données structurées!'. in *World library and information Congres: 78th IFLA general Conference and Assembly*, August 2012, Helsinki, Finland [online] http://conference.ifla.org/sites/default/files/files/papers/wlic2012/80-boulet-fr.pdf (Consulté le 13 juillet 2014)

Bourdon, F. and Boulet, V. (2013) 'VIAF: A hub for a multilingual access to varied collections'. in *World library and information Congres: 78th IFLA general Conference and Assembly*, August 2013, Singapore [online] http://conference.ifla.org/past-wlic/2011/79-bourdon-en.pdf (Consulté le 13 juillet 2014)

IFLA Working Group on FRANAR (2013) *Functional Requirements for Authority Data: a conceptual model (FRAD)* [online] IFLA Working Group on Functional Requirements and Numbering of Authority Records, First edition, 2009 as amended and corrected through July 2013 http://www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf (Consulté le 13 juillet 2014)

IFLA Working Group on FRANAR (2008). *A Review of the Feasibility of an International Standard Data Authority Number (ISADN)* [online] IFLA Working Group on Functional Requirements and Numbering of Authority Records. Available from the page http://archive.ifla.org/VII/d4/wg-franar.htm Direct link http://archive.ifla.org/VII/d4/franar-numbering-paper.pdf (Consulté le 13 juillet 2014)

IFLA (2013). *IFLA Professional Statement on Universal Bibliographic Control*. Initiated by the Bibliography Section, endorsed by the Cataloging Section and the Classification and Indexing Section and finally by the IFLA Professional Committee. http://www.ifla.org/publications/ifla-professional-statement-on-ubc (Consulté le 13 juillet 2014)

Hickey, Th. (2014). 'FRBRizing WorldCat' *Ongoing* [online] 9 April http://outgoing.typepad.com/outgoing/2014/04/frbrizing-worldcat.html (Consulté le 13 juillet 2014)

International Organisation for Standardization (2012) *ISO 27729: 2012: Information and documentation - International standard name identifier (ISNI)*. Geneva, ISO http://www.iso.org/iso/catalogue_detail?csnumber=44292 (Consulté le 13 juillet 2014)

ISNI-IA (2013) '*ISNI for libraries*' [Brochure, documentation provided on the ISNI official website] http://www.isni.org/filedepot_download/126/346 (Consulté le 13 juillet 2014)

LCC (2014a) '*Principles of identification*' Version 1.1 [online] April 2014. http://media.wix.com/ugd/bff7bc_20e274d80da1462ebee53df9f69ab3db.pdf (Consulté le 5 juin 2014)

LCC (2014b) '*The LCC Manifesto*' Version 1.0 [online] April 2014 http://media.wix.com/ugd/bff7bc_a39299633ee74d1e82686c79772c95a6.pdf (Consulté le 5 juin 2014)

Mac Ewan A., Angjeli A. & Gatenby J. (2013) 'The International Standard Name Identifier (ISNI): The Evolving Future of Name Authority Control', *Cataloging and Classification Quarterly*, Vol. 51, No. 1/3, 2013. [abstract online] http://www.cataloging andclassificationquarterly.com/ccq51nr1-3.html (Consulté le 5 juin 2014) http://dx.doi.org/10.1080/01639374.2012.730601

Mac Ewan, A. (2013) '*ISNI, VIAF and NACO and their relationship to ORCID'* [online] Discussion paper for PCC Policy Committee, 4 November. http://www.loc.gov/aba/pcc/documents/ISNI%20PoCo%20discussion%20paper%202013.docx (Consulté le 13 juillet 2014)

Plassard, M.-F. (2003) 'IFLA and Authority Control'. in *Authority control: definition and international experiences: international conference*, 10-12 February 2003. Florence, Italy 2003 http://www.sba.unifi.it/ac/relazioni/plassard_eng.pdf (Consulté le 13 juillet 2014)