

Идентификатори на имена

Милена Миланова

На доцент Нина Шуманова, с уважение и благодарност за подкрепата и доверието.

В света на цифровата информация преодоляването на нееднозначността на данните е от изключителна важност. Използването на уникални идентификатори за книги (ISBN), за продължаващи издания (ISSN), за музикални издания (ISMN), за звукозаписи (ISRC), за аудиовизуални произведения (ISAN) или за музикални произведения (ISWC) отдавна е доказало своето удобство и улеснение в процесите по издирване и управление на информация. Показало е и своите ползи за създателите на съдържание, и за прилагането и осъществяването на различни процеси и дейности по отношение на интелектуална собственост, реализиране на различни бизнес модели и др. В този смисъл правилното установяване на създателя на какъвто и да е обект с интелектуално или художествено съдържание е изключително важно и отдавна познато като проблем за библиотечните специалисти.

Основата, върху която се градят подходите за постигане идентифицирането на имена може да се открие в една от основните и водещи програми на ИФЛА за универсален библиографски контрол, която доказва устойчивостта си през всичките години на своето съществуване и развитие*. Един от резултатите от дейностите по нейното поддържане е свързан с изработването

* Програмата УБК е създадена от ИФЛА през 70-те години на 20 век, като в 2012 г. получава ново развитие чрез приемане на Декларация (<http://www.ifla.org/files/assets/bibliography/Documents/ifla-professional-statement-on-ubc-en.pdf>).

на правила и ръководства за създаване на така наречените контролни файлове, в които се представят установени и вариантни форми на отделните идентичности и връзките между тях. Такива файлове се разработват и поддържат за имена на лица, за наименования на колективни органи, за наименования на анонимни класически произведения, за наименованията на предметни рубрики и др. Основната идея е всяка една страна или регион да има отговорност за изграждането и поддържането на контролни национални файлове, което ще гарантира идентифицирането на обектите по национален, териториален или езиков признак¹.

Дейността на ИФЛА по идентифицирането на имена е свързана с проучване и работа по създаването на уникален номер за имена, която е прекратена поради съвпадение с подобна дейност на Международната организация по стандартизация ISO, която довежда до създаването на Международен стандарт за идентифициране на наименования (ISO 27729:2012 International Standard Name Identifier ISNI²). Дългогодишните усилия на специалистите за гарантиране на контрол върху формите на съществуване на различни идентичности е особено осезаемо и видимо чрез утвърждаването на този стандарт, който всъщност в много голяма степен е повлиян от принципите на нормативния контрол в библиотечната теория и практика³. През 2015 година стандартът е преведен и утвърден на български език като БДС ISO 27729:2015 Международен стандартен идентификатор на наименованията (ISNI)⁴. Този факт подсказва, че като следваща стъпка би трябвало

¹ За по-подробна информация Вж. **Карачожджукова**, Надя. Повишаване качеството на библиотечните онлайн каталози (ОРАС) чрез контролиране на точките за достъп. В: *Литература... Наука... Библиография*: Сб., посветен на 70-год. юбил. на доц. д-р Татяна Янакиева. София: Унив. изд. Св. Климент Охридски, 2014, с. 69-80.

² ISO 27729:2012 Information and documentation – International standard name identifier (ISNI). In: ISO [online]. Online Browsing Platform (OBP). [Viewed on 12.01.2016]. Available from: <https://www.iso.org/obp/ui/#iso:std:iso:27729:ed-1:v1:en>

³ **Angeli**, Anila, **MacEwan**, Andrew, **Boulet**, Vincent. ISNI and VIAF – Transforming ways of trustfully consolidating identities. Paper presented at: IFLA WLIC 2014. In: *IFLA WLIC 2014* [online]: *Libraries, Citizens, Societies: Confluence for Knowledge. Session 86, 16-22 August 2014, Lyon*. Lyon: IFLA, 2014, p. 7. [Viewed on 12.01.2016]. Available from: <http://library.ifla.org/id/eprint/985>

⁴ БДС ISO 27729:2015 Информация и документация. Международен стандартен идентификатор на наименования (ISNI). В: Български институт по стандартизация [онлайн]. [Прегледан на 12.01.2016]. Достъпно от: http://www.bds-bg.org/bg/standard/?natstandard_document_id=72976

да е организирането на дейността по присъждането и въвеждането в употреба на стандартния номер.

Целта на ISNI е да определи стандартен идентификатор на наименования, като те могат да бъдат на физически или на юридически лица и служат за публично представяне на тези лица. Утвърждаването на стандарта е свързано с постигане на унифицираност и избягване на грешки при определяне на наименования, свързани със „създаване, произвеждане, управление и разпространение на медийно съдържание“⁵. Идентификаторът е замислен като „мостов“⁶, който позволява обмен на информация между различни организации, и общности, с цел нееднозначно идентифициране на наименования (лица, колективи, групи) изявяващи се в различни творчески среди. Информацията, записвана в един запис е достатъчна за идентифицирането на дадено наименование без да нарушава права, свързани с интелектуална собственост или защитата на лични данни. Номерът е съвместим с всички стандартни идентификатори, разработени от ISO. Представлява глобален идентификатор от висок клас и позволява осъществяването на високо ниво на автоматизация и управление на съвместима информация за лица и организации в цифрова среда. Той не само улеснява откриването на информация, но и рационализира бизнеса, независим е от езици и системи, както и от различни сфери и географски територии. Това всъщност го определя като мостов идентификатор или като хъб между множество сфери - свързан е с контролните национални файлове, но преодолява тяхното ограничение по отношение на обектите и национален, езиков или териториален белег.

ISNI определя обектите като публични идентичности на страни. Като под публична идентичност се разбира „идентичност на дадена страна или псевдоним, с която тя е представена публично“, а под страна се разбира „физическо или юридическо лице (независимо дали е официално регистрирано) или група от такива лица“⁷.

Един номер се присъжда само на една публична идентичност, т.е. ако едно и също лице е публично известно с официално име и с псевдоним, то всяка идентичност (и името и псевдонима) ще получат отделен идентификационен номер, като вариантите

⁵ Пак там, с. 6

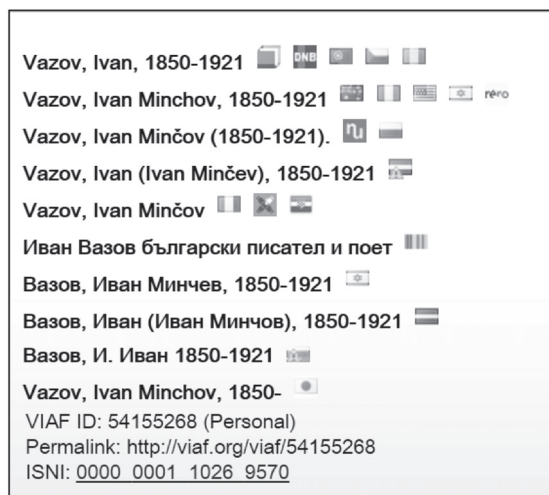
⁶ Пак там, с. 16

⁷ БДС ISO 27729:2015 ..., с. 6

на една идентичност представени чрез различни начини на изписване, алтернативно представяне, различни азбуки и езици не определят присъждането на отделни номера. Промените в наименованието на юридическо лице при сливане или разделяне изисква присъждането на отделен ISNI. Номер е възможно да се присъжда и на групи от физически или юридически лица. Важно е да отбележим, че номерата се присъждат само веднъж не могат да бъдат използвани отново.

При регистрацията и присъждането на номер на всяка една публична идентичност е необходимо да се попълнят определен набор от метаданни, които включват наименованието на тази идентичност, вида на страната (лицето) и поне един от елементите връзка към външни данни, творческа категория или роля. Регистрацията трябва да се извършва от специализиран орган, който определя правилата и наборът от метаданни към идентичностите, които искат да бъдат регистрирани.

Идентификаторът се състои от 16 цифри, които представляват символичен номер без никакво смислово натоварване. Изписва се в четири блока от по четири цифри, като последната е контролна. Предхожда се от абревиатурата ISNI. Така например към публичната идентичност Вазов, Иван Минчов е с присъден идентификатор *ISNI 0000 0001 1026 9570*⁸ (изображение. 1).



Изображение 1. Пример за публична идентичност на физическо лице с идентификатор ISNI

⁸ **ISNI** [online]. *ISNI Search*. [Viewed on 13.01.2016]. Available from: <http://www.isni.org/search>

В запис са представени следните елементи:

- вид на страницата (лицето), в което се посочва дали лицето е физическо, юридическо или група;
- име на идентичността, което включва всички видове представяне на едно име;
- дати, свързани с идентичността. Формата на изписване на датата трябва да бъде предварително определена;
- творческа категория или вид творчество, което се изразява чрез предварително определени стойности;
- роля на лицето. Изразява се чрез предварително установени стойности;
- връзки към външни данни (като имена и заглавия) където може да се намери допълнителна информация за публичната идентичност. Този елемент може да съдържа различни стойности в зависимост от регистрираната идентичност. Видът на връзките и допустимите стойности трябва да са предварително определени.

Към записите могат да бъдат представени и други елементи от метаданни като:

- място и вид на мястото, с предварително определена форма и стойност на данните;
- свързан идентификатор, който е определен като ISNI на друга публична идентичност, например при псевдоним⁹.

Структурата на записите в базата на системата ISNI се доближава изключително до създаването в библиотеките контролни файлове. Чрез асоциирането на номер се постига едноточност на отделните идентичности. Концепцията за мостов идентификатор подпомага обменът на информация между различните творчески индустрии, които поддържат различни набори от метаданни за идентифицираните от тях идентичности. В тези условия номерът ISNI чрез поддържането на минимален набор от метаданни, достатъчни за разграничаване на две публични идентичности позволява да се обменя информация между партньори без да се нарушават правата за достъп до защитена информация. Или ISNI опосредства диалога между отделни бази от данни като установява отделните публични идентичности и тяхната свързаност.

⁹ БДС ISO 27729:2015 ..., с. 13-16.

Базата от данни на ISN* съдържа записи за 8,99 милиона публични идентичности, включващи 8,46 милиона лица (от тях 2,55 милиона изследователи) и 525,636 организации. ISNI е обединяваща база, на ресурсите на 30 институции и бази данни и 40 големи национални и изследователски библиотеки. В един ресурс се обединяват публичните идентичности на изследователи, изобретатели, автори, творци в сферата на изкуството, изпълнители, издатели и др. Чрез присъждането на уникален номер към всяка отделна идентичност се постига разрешаване на проблемите при търсене и откриване на идентичности и тяхната нееднозначна свързаност с всяка публикувана творба и нейния създател в момента, в който тази творба бъде описана някъде.

Създаването на централната база данни на ISNI е иницирано през 2011 г., въз основа на библиотечните контролни файлове, чрез данните във виртуалния международен контролен файл VIAF¹⁰. Значителна част от идентичностите във VIAF имат асоцииран номер ISNI. Метаданните в записите на идентичностите в ISNI са извлечени от данните в контролния файл.

Основната разлика в създаването на двете бази от данни е в подхода към обектите. VIAF се фокусира върху нормативния контрол на библиографски обекти (лице, фамилия, колективен орган, произведение, проявление, събитие и място*) като в зависимост от съответните национални правила за каталогизация библиографските обекти могат да бъдат възприемани различно¹¹. ISNI от своя страна се концентрира върху идентифицирането на публичната идентичност на лица, колективни органи, групи като не изгражда задължителна връзка между индивида и неговата публична идентичност, освен ако това не е изрично позволено от отделния индивид или колективен орган¹².

* Данните са към април 2016 г. Вж. **ISNI** [online]. Available from: <http://www.isni.org/>

¹⁰ **VIAF** [online] *The Virtual International Authority File*. Available from: <http://viaf.org/>

* Според концептуалния модел за функционалност на библиографските и контролни записи.

¹¹ **FRANAR** [онлайн]. Функционални изисквания към нормативните данни. Концептуален модел. ИФЛА, 2009, с. 10 [Прегледан на 14.01.2016]. Достъпно от: http://www.ifla.org/files/assets/cataloguing/frad/frad_2009-bg.pdf

¹² **Angjeli, Anila, Mac Ewan, Andrew, Boulet, Vincent.** ..., p. 3.

За българските специалисти нормативният контрол на имената на лица е познат и постепенно намира място все по-активно в процесите по каталогизация. Международният контролен файл на имена е изграден на принципа на създаването на обединена база с нормативна информация. Обединената база постепенно се разраства и включва все повече национални файлове с нормативна информация, както и обединени ресурси на различни организации като Wikidata¹³, Сводният списък на имената на представители на изкуството (ULAN)¹⁴ и др. с фокус върху различни специфични области. Основните цели са свързани с споделяне на данни, намаляване на разходи, осигуряване на контролни данни в различни формати (RDF, MARCXML, JSON), езици и писмени знаци, от които потребителите имат нужда. Крайните потребители могат да използват данните свободно. Към началото на 2016 г. базата данни съдържа около 30 000 000 записа¹⁵.

VIAF не създава никакви данни сам по себе си, а само комбинира и обединява данни доставени от членовете му. Важно е да се отбележи, че базата е неутрална по отношение на политиката на каталогизиране на своите партньори и не се занимава с използваните от тях работни формати. Файлът има отношение към разнообразни модели от данни, структура на данните и стандарти за каталогизация. Изгражда мостове, без да налага правила или добри практики на своите партньори. При обединяването на информацията, подавана от различните участници се търси съответствие между данните. Използва се напълно автоматизиран алгоритъм, който обединява данните на различните доставчици в общи записи или клъстери за всеки отделен обект. Данните се съпоставят и групират без никаква ръчна намеса и без каквото и да е отношение към качеството им, което се гарантира от участниците в системата. При необходимост, поява на нов участник или актуализиране на данните в базите на партньорите, алгоритъмът отново проверява обединените данни и ако е необходимо ги коригира според новите условия. В базата данни на VIAF също се присъждат идентификационни номера, като те се присвояват към определен клъстер от данни¹⁶.

¹³ **Wikidata** [online]. Available from: https://www.wikidata.org/wiki/Wikidata:Main_Page

¹⁴ **Union List of Artist Names** (Getty Research Institute) [online]. Available from: <http://www.getty.edu/>

¹⁵ **VIAF** ...

¹⁶ **Angjeli, Anila, Mac Ewan, Andrew, Boulet, Vincent. ...**, p. 7.

Идентификационният номер на клъстер за обект лице Вазов, Иван Минчов присвоен от VIAF е ID: 54155268¹⁷ (изображение 2.).



Изображение. 2. Пример за клъстер за обект физическо лице с идентификационен номер от VIAF

Клъстерът обединява данните от 20 контролни файла, като представя различни установени форми на името и 104 алтернативни форми, включително и данни от Wikidata.

Поддържането и управлението на двете системи е тясно свързано не само поради първоначалното изграждане на базата на ISNI въз основа на данните във VIAF. Анализът на Прегледаните данни и отчетите се подготвят от екип на OCLC ISNI в Лайден (Холандия), който се явява ISNI Агенция по прехвърляне (Assignment Agency). Оценката на качеството на базата данни въз основа на текуща база се осъществява от експерти по нормативен контрол от Френската национална библиотека и Националната библиотека на Великобритания, които образуват екипа за качество на ISNI. Към международния контролен файл е изграден Съвет на VIAF, в който е представен всеки отделен доставчик на

¹⁷ VIAF [online] *The Virtual International Authority File. Search*. [Viewed on 14.01.2016]. Available from: https://viaf.org/viaf/54155268/#Vazov,_Ivan_Minchov,_1850-1921

данни и който дава препоръки към OCLC относно подобряването на базата данни¹⁸.

Всъщност развитието и на двете системи е паралелно и взаимно се допълва като отдава полагащото се място на дейността на библиотечните специалисти по отношение унифициране на информацията. Информатизацията на библиографската дейност, нарастването на информационните потоци, появата на нови, по-бързи канали за създаване и публикуване на съдържание в уеб пространството водят до решения, които показват и друг подход: от споделяне към контрол. В контекста на Отворените свързани данни (Linked Open Data) непрекъсваното нарастване на обмена на данни, проблемът за качеството на данните и тяхната достоверност се превръща във все по-стратегически и основен. Необходимо е да се гарантира надеждността на свързаността на данните в мрежата. Допускането на неточности ще доведе до погрешното идентифициране на данните и грешки при изграждане свързване на информацията. Можем да твърдим, че изработването и поддържането на контролни файлове за различни обекти и идентичности и споделянето на тази информация е от изключителна важност за съвременния свят, който в голяма степен е зависим от наличието и достъпа до бърза, качествена и проверена информация.

¹⁸ Angjeli, Anila, Mac Ewan, Andrew, Boulet, Vincent. ..., pp. 9-10.