



FAIR

ПРИНЦИПИ

ПРИЛОЖИМОСТ ЗА
НАУЧНИ ИНФРАСТРУКТУРИ И ДИГИТАЛНИ
ХРАНИЛИЩА ЗА КУЛТУРНО СЪДЪРЖАНИЕ

Калина Сотирова-Вълкова
Проект Кла-ДА

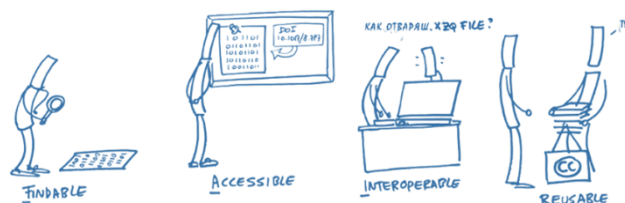
юли, 2020

СЪДЪРЖАНИЕ

НАЧАЛОТО	2
СЪЩНОСТ	3
КАКВО ПРАВИ ЕДНА УСЛУГА/СИСТЕМА FAIR?	4
КАКВО ОЗНАЧАВА ТОВА ЗА ИЗСЛЕДОВАТЕЛИТЕ?	4
КАКВО ОЗНАЧАВА ТОВА ЗА ИНСТИТУЦИИТЕ НА ПАМЕТТА?	5
FAIR ПРИНЦИПИ И ПОСТОЯННИ ИДЕНТИФИКАТОРИ	5
FAIR: ТЕХНИЧЕСКА ЕКОСИСТЕМА.....	7
ДОБРИ ПРАКТИКИ И ДОКАЗАНИ С РЕЗУЛТАТИ FAIR ИНСТРУМЕНТИ.....	11
Use Case 1. HERITAGE CONNECTOR – КЪМ НАЦИОНАЛНА КОЛЕКЦИЯ?.....	12
Use Case 2. НАЦИОНАЛНАТА ГАЛЕРИЯ ВЕЛИКОБРИТАНИЯ: FAIR + CIDOC-CRM & LOD.....	13
ИЗТОЧНИЦИ	17
ПРИЛОЖЕНИЕ. IIIF – ИЗОБРАЖЕНИЯ В НАУЧНИ ИНФРАСТРУКТУРИ.....	18

FAIR ПРИНЦИПИ

УПОТРЕБИ И ПРИЛОЖИМОСТ КЪМ ОБЕКТИ НА КУЛТУРНОТО НАСЛЕДСТВО В ДИГИТАЛНИ ХРАНИЛИЩА И НАУЧНИ ИНФРАСТРУКТУРИ¹



НАЧАЛОТО

Появата на инициативата FAIR² през 2016 г. се базира на нуждата от добро управление на разнородни данни, както и подобряване на функционалностите на инфраструктурата, подкрепяща повторната употреба на *научни данни*, нужда осъзнавана както от академични среди, индустрия, финансиращи агенции, но и от институциите на паметта, и научни е-инфраструктури. Още в първите формални публикации по темата, *The FAIR Guiding Principles for scientific data management and stewardship*³, от март 2016, появила се успоредно със сайта [<https://www.go-fair.org>], се декларира, че съществуването на FAIR-принципите зависи от наличието и устойчивостта на услугите, които подпомагат създаването, (повторното) използване и запазване на резултатите от изследователската дейност:

*“This article describes four foundational principles — Findability, Accessibility, Interoperability, and Reusability—that serve to guide data producers and publishers as they navigate around these obstacles, thereby helping to maximize the added-value gained by contemporary, formal scholarly digital publishing. Importantly, it is our intent that the principles apply not only to ‘data’ in the conventional sense, but also to the algorithms, tools, and workflows that led to that data. All scholarly digital research objects—from data to analytical pipelines—benefit from application of these principles, since all components of the research process must be available to ensure transparency, reproducibility, and reusability.”*⁴

¹ Авторът предлага този текст за вътрешна употреба на координаторите на проекта КЛАДА-БГ. Цитираните ресурси съзнателно са оставени на оригиналния английски език, като превод е направен, където се налага.

² Инициатива [<https://www.go-fair.org>] Посл. посетен: 10-07-2020

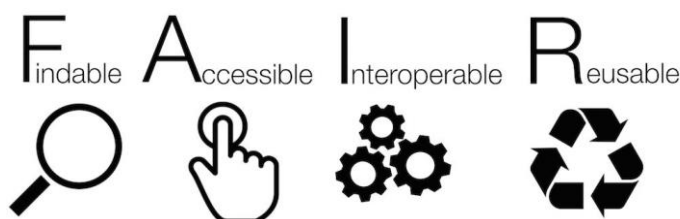
³ [<https://doi.org/10.1038/sdata.2016.18>] Посл. посетен: 10-07-2020

⁴ Пак там, с. 3

СЪЩНОСТ

Основните принципи на FAIR са разписани в ключови документи на едноименната фондация [<https://doi.org/10.1038/sdata.2016.18>] и могат да бъдат резюмирани така:

1. *Откриваеми:* **(O1)** (meta)data are assigned a globally unique and persistent identifier, **(O2)** data are described with rich metadata (дефинирани от П1 по-долу), **(O3)** metadata clearly and explicitly include the identifier of the data it describes, **(O4)** (meta)data are registered or indexed in a searchable resource
2. *Достъпни:* **(D1)** (meta)data are retrievable by their identifier using a standardized communications protocol, **(D2)** the protocol is free, open and universally implementable, **(D3)** the protocol allows for an authentication and authorization procedure, where necessary, **(D4)** metadata are accessible, even when the data are no longer available
3. *Оперативно съвместими:* **(C1)** (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation, **(C2)** (meta)data uses vocabularies that follow FAIR principles, **(C3)** (meta)data include qualified references to other (meta)data
4. *Повторно използваеми:* **(P1)** (мета) данните са детайлно описани с множество точни и подходящи атрибути, **(P2)** (мета) данните се представят с ясен и достъпен лиценз за използване, **(P3)** (мета) данните са свързани с генерирането на данни, **(P4)** (мета) данните отговарят на приложимите за общността стандарти



Четири основни принципа на FAIR поставят специфичен акцент върху повишаването на способността на машините автоматично да намират и

използват данните, в допълнение към подпомагането на повторната им употреба от хора. Сравнително новопоявила се идея е, тези принципи, с постоянните идентификатори като техни инструменти, да се прилагат и към обекти на КИН⁵, за да ги описват разбираемо и да улесняват изграждането на LOD и граф на знанието. Основните ползи са за премахване на двусмислието, за интердисциплинарни изследвания. Това е възможно в е-инфраструктура, която има механизмите за дългосрочна поддръжка на споменатите услуги.

⁵ HERITAGE PIDs - <https://twitter.com/HeritagePIDs> Посл. посетен: 10-07-2020

КАКВО ПРАВИ ЕДНА УСЛУГА/СИСТЕМА FAIR?

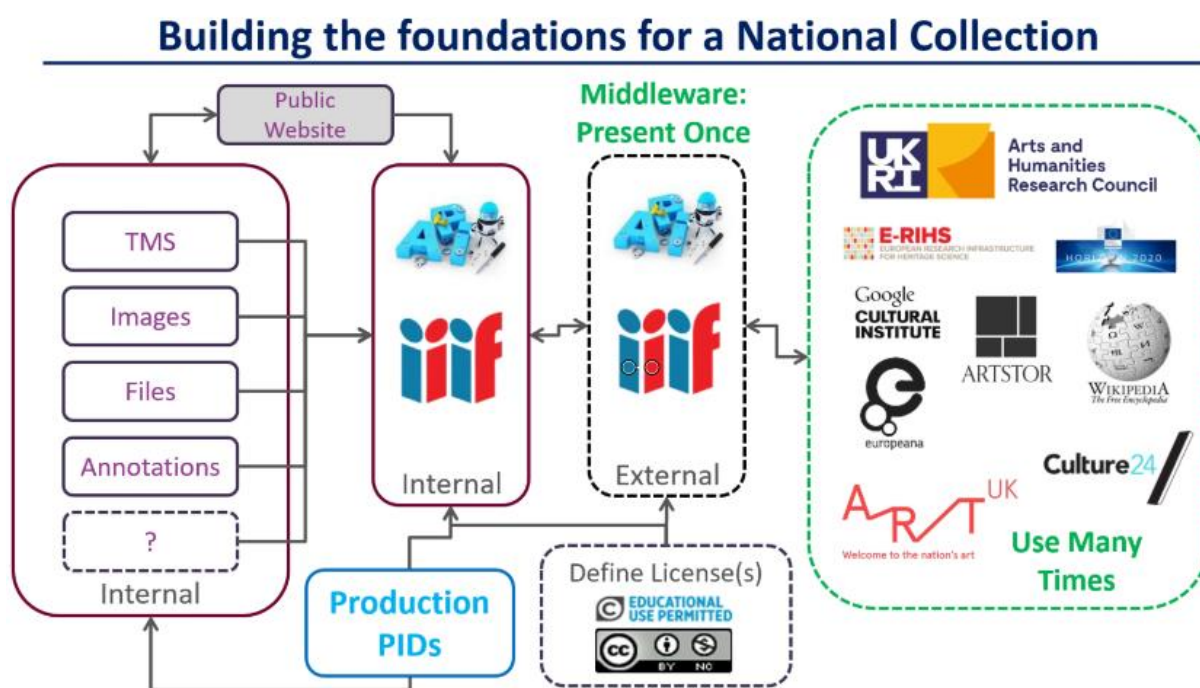
Четири основни принципа (откриваеми, достъпни, оперативно съвместими и повторно използваеми обекти) предполагат добре описани данни да се идентифицират, структурират и контекстуализират за различни видове употреба и различни потребители. Това става чрез инструменти като постоянните идентификатори, сертифициране на хранилищата, услуги за управление и дългосрочно поддържане на данните в тези хранилища и др. На инструментите, сертификати, специализирани услуги и идентификатори на обекти ще се обърне специално внимание в този текст, с фокус върху евентуалното им приложение в научната инфраструктура КЛАДА-БГ, в която към юли 2020г има информация от обекти на книжовната памет и етнографското наследство, с тенденция към разширяване с научни публикации и изследвания за безписменото тракийско наследство и археологията. *Докато първите два принципа, откриваемост и достъп*, изискващи по скоро техническа и технологична експертиза се адресират от основни услуги като каталогизиране, постоянни идентификатори (PID) и storage, то *вторите два принципа, повторната употреба и оперативната съвместимост*, изискват развитие на услуги, свързани със специфичната област на приложение и употреба на ресурсите в конкретно хранилище: онтологии, общи и тематични, тезауруси, подредба (curation), два типа стандартизирани метаданни (generic и domain specific), стопанисване (stewardship). Накратко, FAIR стандартизира споделянето на данни, така че неунифицирано, непоследователно каталогизирани данни, публикации, и обекти от музейни колекции да са лесно откриваеми и използваеми.

КАКВО ОЗНАЧАВА ТОВА ЗА ИЗСЛЕДОВАТЕЛИТЕ?

Дистрибутираните заявки, управлявани от софтуер за посредничество (brokering software) се ползват за виртуално интегриране на данни. Необходимостта от *разпределен анализ* за различни набори от данни (datasets) е един от основните двигатели за използване на FAIR данни: той изисква метаданни за намиране на ресурсите, протоколи за достъп до тях, съгласувани спецификации, така че данните да са оперативно съвместими и пре-използвани с доверие. Чрез достъпа до хранилища, базирани на FAIR принципите, изследователите ще имат възможност да повишат качеството в изследователската си дейност, да правят интердисциплинарни връзки между изследваните обекти, да ги поставят в различни контексти, и най-вече генерираните публикации да бъдат бързо споделяни, повторно ползвани и уникално идентифицирани.

КАКВО ОЗНАЧАВА ТОВА ЗА ИНСТИТУЦИИТЕ НА ПАМЕТТА? FAIR ПРИНЦИПИ И ПОСТОЯННИ ИДЕНТИФИКАТОРИ

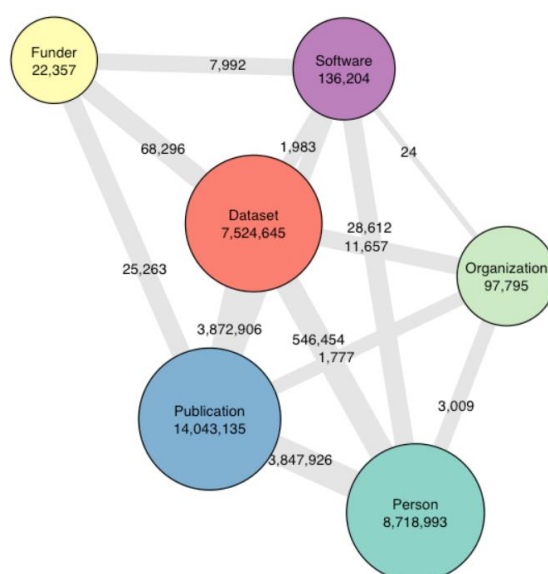
Докато проблема е донякъде решен когато става дума за научни публикации, вкл.чрез портали като OPEN AIRE+, където България е партньор, и European Open Science Cloud (EOSC), за обекти на КИН това не е така. Тези обекти трябва също да бъдат идентифицирани в научни инфраструктури, и още по-важно ситуирани в контекст, и свързани с други мрежи, за да са приложими и към тях FAIR принципите. Постоянният идентификатор осигурява стабилен и дългосрочен линк към описан обект (entity described) или информация за него. Важното при взимане на информирано решение за използването му в институции на паметта е да се знае предварително, че дългосрочността и устойчивостта на използването му не зависи от технологията, а от управлението. PID може да служи като базова инфраструктура за изграждане на национална колекция (Фиг. 1) чрез подход, който обединява разнородни колекции и научни дисциплини.



Фиг.1. Модел на Великобритания за създаване на национална колекция

В този смисъл е интересно да се спомене как се развива графът за разпространението и използването на постоянните идентификатори в световен мащаб – фиг. 2. Графът от фигура 2 включва всички DOI номера на DataCite, 9 млн. Crossref DOIs, всички ID номера на ORCID, и всички Research Organization Registry (ROR), Crossref Funder ID, and Registry of Research Data Repositories (re3data) records, for a total of about 35 million resources with PIDs and associated metadata, and about 9 million links between them.

PID Graph
Number of nodes and connections (08 Mai 2020)



Фиг.1 [PID Graph](#) – 8 май 2020

Проектът **FREYA**⁶, който актуализира този граф 24/7 представлява инфраструктура за PID се развива на три основни стълба: PID граф, PID форум и PID Commons, чието разглеждане може да бъде от особена полза за екипа на КЛАДА-БГ с оглед възможни сценарии за прилагане на постоянните идентификатори у нас. Такъв сценарий за държавната научна политика, и за отделните изследователски звена на национално ниво би означавало следване на дългосрочна и устойчива във времето политика за публикуване в издания, които имат платен и поддържан DOI регистър. Важно условие за реализация на такава политика е познаването на механизма на работа на доставчиците на постоянни идентификатори в световен мащаб. Трите най-авторитетни **Persistent Identifier** доставчици са: **DataCite**, **CrossRef**, **ORCID**, като за научна информация се препоръчва предоставя постоянни идентификатори (DOI) за изследователски данни и други резултати от научните изследвания.

⁶ FREYA проект [https://www.project-freya.eu/en/about/mission] Посл. посетен: 10-07-2020

Организациите в рамките на изследователската общност се присъединяват към DataCite като членове, за да могат да присвояват DOI на всички свои изследователски резултати. По този начин техните резултати стават откриваеми и свързаните метаданни се предоставят на общността.

International Standard Name Identifier (ISO 27729), част от семейството международни стандартни идентификатори включва идентификатори за артефакти, работи, записи, продукти и носители на права във всички репертоари, напр. **DOI, ISAN, ISBN, ISRC, ISSN, ISTC, и ISWC**. Той се използва от носители на права и творчески организации (изследователи, откриватели, писатели, артисти, дизайнери, издатели, агрегатори и др.): Мисията на **ISNI International Authority (ISNI-IA)** и да свърже с публичното име на издателя, автора, изследователя, откривателя, писателя, артиста уникален постоянен идентифициращ номер, за да елиминира двузначността в имената; и да разпространи всеки ISNI номер в глобалната мрежа за доставка, вкл. търговска, като целта е всяка публикуван цифров обект да бъде недвусмислено свързан точно и само със своя автор/носителя на правата, в което и хранилище да е описан този обект. ISNI по този начин се явява като мост, използваем от различни области, научни, приложни, вкл. индустрии, критично важен за Linked Data и уеб приложенията на Семантичната мрежа.

FAIR: ТЕХНИЧЕСКА ЕКОСИСТЕМА

Централно значение за реализацията на FAIR са **цифровите обекти**: данни, софтуер, протоколи или други изследователски ресурси, които трябва да бъдат придружени от постоянни идентификатори (PIDs) и метаданни достатъчно богати, за да може надеждно да бъдат намерени, използвани и цитирани. Освен това данните трябва да бъдат представени в идеалния случай с отворени формати и да са описани детайлно чрез стандарти за метаданни и речници. Стандартите, речниците, форматите трябва да са приети от свързаната изследователска общност, за да се даде възможност за оперативна съвместимост и повторна употреба. Софтуерът и алгоритмите, когато се споделят, трябва да включват не само самия източник, но и машинно изпълними изявления за лицензиране.

За прилагането на основната концепция за **FAIR цифров обект**, две области от дейности са особено важни: 1) разработване, усъвършенстване и приемане на споделени речници, онтологии, спецификации на метаданни и стандарти за осигуряване на оперативната съвместимост и повторна употреба, което изисква по-координирани усилия, и 2) осигуряване на дългосрочно и стабилно управление на данните, хранилищата и услугите, което изисква

професионално ангажиране на изследователски инфраструктури, хранилища и самата общност. Хранилищата на данни и услугите следва да предоставят дългосрочно управление на данни с оглед постигане на сертифициране – в този смисъл и за КЛАДА-БГ, CoreTrustSeal следва да се изследва като отправна точка.

Цифровите обекти на FAIR са разположени в по-широка **екологична система FAIR**, включваща услуги и инфраструктура. Развитието на техническата екосистема за FAIR може да бъде решено с комбинация от двата подхода: отгоре-надолу (архитектурни) и отдолу-нагоре (органични, базирани на спецификации). Реализацията на FAIR трябва да включва като минимум пет основни компонента: (1) политики, (2) идентификатори, (3) План за управление на данните, (4) стандарти и (5) хранилища / поддържаща инфраструктура.

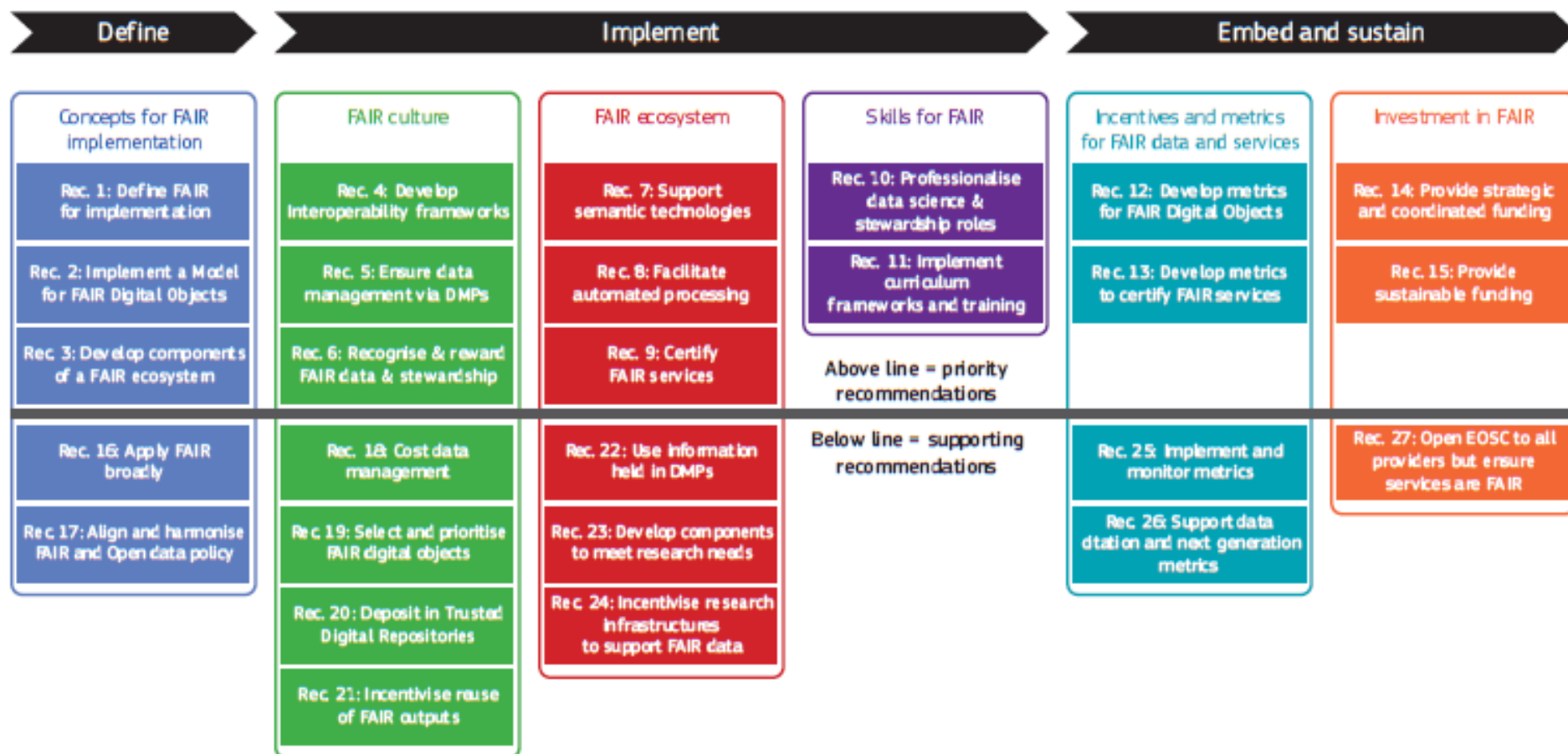
1. **Политики.** В тази екосистема политиките за данни се издават от няколко заинтересовани страни и помагат да се дефинират и регулират изискванията за функционирането на услугите за данни.
2. **Идентификатори.** Постоянните идентификатори са ключов елемент от екосистемата, отговорни за данни, софтуер, институции, изследователи, финансиращи, проекти и инструменти. Характеристики на постоянните идентификатори: Активни (actionable); Globally unique, не може да се използва за друг обект, използва се за disambiguation; Persistent, което с гарантира от комбинация от механизми, административни и управленски, не само технологични; Пр. <https://doi.org/10.17863/CAM.4926>; Действат независимо от URL-а на обекта, като дигиталния обект може да бъде цяла колекция, или само един обект; Enable linked data / спомагат за избягване на двузначност; Поддържат цитиране, което е особено важно за изследвания; Налице е общо разбиране на техните функции сред потребителя.
3. **План.** Плановите за управление на данни предоставят динамичен индекс, който представя информацията, свързана с даден проект и я свързва с различните компоненти на FAIR.
4. **Стандарти.** Спецификациите и стандартите за метаданни, речниците и онтологиите за описание на данни са важни за прехвърляне и обмен на протоколи за достъп до данни, регулиращи сертифицирането на хранилища.
5. **Хранилища.** Репозиториите предлагат бази данни и услуги за данни и трябва да бъдат сертифицирани⁷, за да се гарантира доверието.

⁷ CoreTrustSeal Посл. посетен: 10-07-2020

За работеща и сигурна екосистема са нужни регистри, които да каталогизират компонентните услуги и автоматизираните работни процеси (workflows). Подобна стабилна екосистема задължително ще бъде широко разпространена. Това ще изисква технически механизми за свързване на ресурси, както и механизми за координация и постигане на съгласие относно спецификации и стандарти. Според EK⁸ European Open Science Cloud (EOSC) ще играе важна роля във всеки от тези механизми. Тъй като данните все повече ще се съхраняват на различни места по различни причини (разходи за копирането им, законови или етични ограничения), е необходима процедура за за сключване на споразумения между хранилища или регистри, за да се изпълнят определени задачи съвместно. Това пряко касае и инфраструктурата КЛАДА-БГ, ако нейна цел е превръщане в работеща разпределена FAIR екосистема. (вж. Фиг. 3).

Споделеният опит на Националната галерия на Великобритания, и партньорите по проекта HERITAGE CONNECTOR, описан по-долу, в отговор на въпроса КАК да направим своите данни достъпни, се свежда до три ключови стъпки: (1) опростен регистър за постоянни идентификатори, (2) идентифициране на наличните данни: обекти от колекцията, техни дигитални образи в съответните формати и качество, стандартизирани описания, вкл. за автора и носителя на правата, и (3) построяване и документиране на отворено API, resolvable URIs, използвайки CIDOC CRM за „мапване“ на данните и IIIF за представяне на образите.

⁸ EU Final Report: Turning FAIR into reality, p.12. 2018 [<https://op.europa.eu/s/n9BX>] Посл. посетен: 10-07-2020



Index to FAIR Action Plan recommendations

Priority recommendations sit above the dotted line. These provide key concepts for FAIR Digital Objects and a FAIR ecosystem, which is then implemented through changes in research culture, technology and skills. Metrics, incentives and investment will drive culture change and implementation.

Фиг.3. FAIR ПЛАН ЗА ДЕЙСТВИЕ НА ЕК С ПРЕПОРЪКИ „TURNING FAIR INTO REALITY“ [HTTPS://OP.EUROPA.EU/S/N9BX], СТР. 17

Следват добри практики на реализиращи се в момента проекти свързани с представяне на уникално идентифицирани КИН обекти. Част от тях са извлечени от портала на Европейска Изследователска Инфраструктура за Науки за Наследството, в която консорциумът КЛАДА-БГ може да стане част с оглед устойчивост на резултатите. Чрез достъпа до четирите платформи, създадени по проекта на ЕК, Европейска Изследователска Инфраструктура за Науки за Наследството (ARCHLAB, DIGILAB, FIXLAB, MOLAB), E-RHS поддържа множество екипи с голямо разнообразие от изследвания, от тясно фокусирани през широкообхватни до насочени към дългосрочно сътрудничество проекти. От полза за КЛАДА-БГ би било евентуално включване в [European Research Infrastructure for Heritage Science](#). От [картата](#) на сайта на проекта с националните представителства на инфраструктурата се вижда, че държавите с такива хъбове са: Белгия, Кипър, Дания, Франция, Германия, Гърция, Унгария, Италия, Нидерландия, Полша, Португалия, Румъния, Испания, Словения, Швеция и Обединеното кралство.

ДОБРИ ПРАКТИКИ И ДОКАЗАНИ С РЕЗУЛТАТИ FAIR ИНСТРУМЕНТИ

Белгия. Националната политика на Белгия създава добър модел на работа за успешно прилагане на FAIR принципите чрез множество проекти⁹, от които важни за целите на КЛАДА-БГ биха били: (1) [DCAT-BE](#) (Unify Belgian datasets), (2) The [DataTank](#) (Transforming data formats into web readable formats) и (3) W4P - Crowdsourcing open innovation: отворена платформа за иновативни проекти в областта, която инициира набирането на средства.

Великобритания. Опитът и решението на Великобритания, което към май 2020 се ползва от Националната им галерия и други институции на паметта е *Collections Information Integration Middleware* ([CIIM](#)), предназначен за GLAM институции и за всяка организация, която желае да сподели, разпространи, синхронизира и визуализира свое съдържание описан така: „CIIM свързва хората със съдържанието и може да се ползва за да: публикува автоматично онлайн данните си за колекции (архив, обект, медия и авторитет); предоставя данни на Europeana или Archives Portal Europe; публикуване на данни в triple и ги съотнася към (*align*) с CIDOC CRM; синхронизира DAMS и данни от колекции, за да премахне дублираните данни; добавя DAMS медийни връзки в базата данни за колекции; **използва IIIF** както за представяне на изображения, така и за представяне на метаданни; обработка данни от web analytics, за да представи ползването на данни и услуги“. От интерес за техническия екип на КЛАДА-БГ би била документацията за реализацията на два от CIIM, базираните проекти: [IMUSEUM](#), [SCIENCE MUSEUM](#).

⁹ [<https://be.okfn.org/projects/>] Посл. посетен: 10-07-2020

На портала на КЛАДА-БГ се предвиждат да се качат изображения от музеите и институциите –партньори с различен характер: част от публикация с необявен източник, част от сайта на доставчика, изображения от Интернет с КК лиценз, собствени изображения със защитено авторско право. Как всички те да могат да се използват некомерсиално, по унифициран начин да се преглеждат, да се аотират и пре-използват за целите на проекта? Възможно решение е ползването на посочената по-горе **International Image Interoperability Framework** [<https://iiif.io>]. IIIF възниква в резултат от усилията на разрастваща се общност на водещи световни изследователски библиотеки и хранилища за изображения за съвместно сътрудничество по създаване на оперативно съвместима технология и обща рамка за достъпване до дигитални образи¹⁰. **Повече за IIIF вж. в Приложението на с. 16**

Според разработчиците на друга подобна платформа, **ZENODO**, ползвана от много европейски университетски библиотеки и CERN, решението на основното предизвикателство пред обединението на разнородните данни и приложението на FAIR принципите е: *Map metadata to DataCite with detail loss ; Simplified custom metadata (searchable): Key-value pairs based on vocabulary (JSON-LD)] <http://rs.tdwg.org/dwc/terms/collectionCode>] Domain-specific metadata files (not searchable, като напр. DarwinCore XML*

Use Case 1. HERITAGE CONNECTOR – КЪМ НАЦИОНАЛНА КОЛЕКЦИЯ?

HERITAGE CONNECTOR е текущ проект на University of London, Science Museum Group и няколко разнородни музея във Великобритания (вкл. изкуство, етнография и дейности по дигитализация) целящ трансформиране на текст в данни за извличане на знания и създаване на семантични връзки. Чрез HERITAGE CONNECTOR седем институции във Великобритания ползват постоянни идентификатори на дигитални обекти от КИИ и нещо повече, обединяват разнородни данни от различни колекции в една обща „национална колекция“. Целта на проекта има много общо с целите на националната научна инфраструктура Клада-БГ, поради което технологичните решения биха могли да послужат като работеща добра практика, от която да се заимства. Партньорите по проекта са: British Library, Royal Botanic Garden Edinburgh (Lorna Mitchel), Natural History Museum (Matt Woodburn), National Gallery (Joseph Padfield), University of Glasgow (Rod Padge), ScienceMuseumShop (Jack Kirby), .V&A museum (Richard Palmer).

¹⁰ Демо и приложения на IIIF [<https://iiif.io/apps-demos/>] Посл. посетен: 10-07-2020

PID Details

This basic search system has been setup to display the current details associated with a given PID.

Details for: 0F6J-0001-0000-0000 - raw {JSON}

```
{
  "identifier": {
    "type": "object number",
    "object_number": "NG35",
    "value": "NG35",
    "primary": true
  },
  "pid_tms": "0F6J-0001-0000-0000",
  "type": "pid_tms",
  "value": "0F6J-0001-0000-0000",
  "display_number": "NG35",
  "type": "display number",
  "value": "NG35",
  "sort_number": "NG 35",
  "type": "sort number",
  "value": "NG 35",
  "@link": {},
  "PID_(NG_alternative)": "000-033T-0000",
  "type": "PID (No alternative)",
  "value": "000-033T-0000"
}
```

Production Local PIDs:

- Create new, independent IDs for all “things” one wants to discuss or describe.
- namespace/AAAA-AAAA-AAAA-AAAA
 - <https://data.ng-london.org.uk/0F6J-0001-0000-0000> -> Details of NG35

Consider the process of globally registered PIDs:

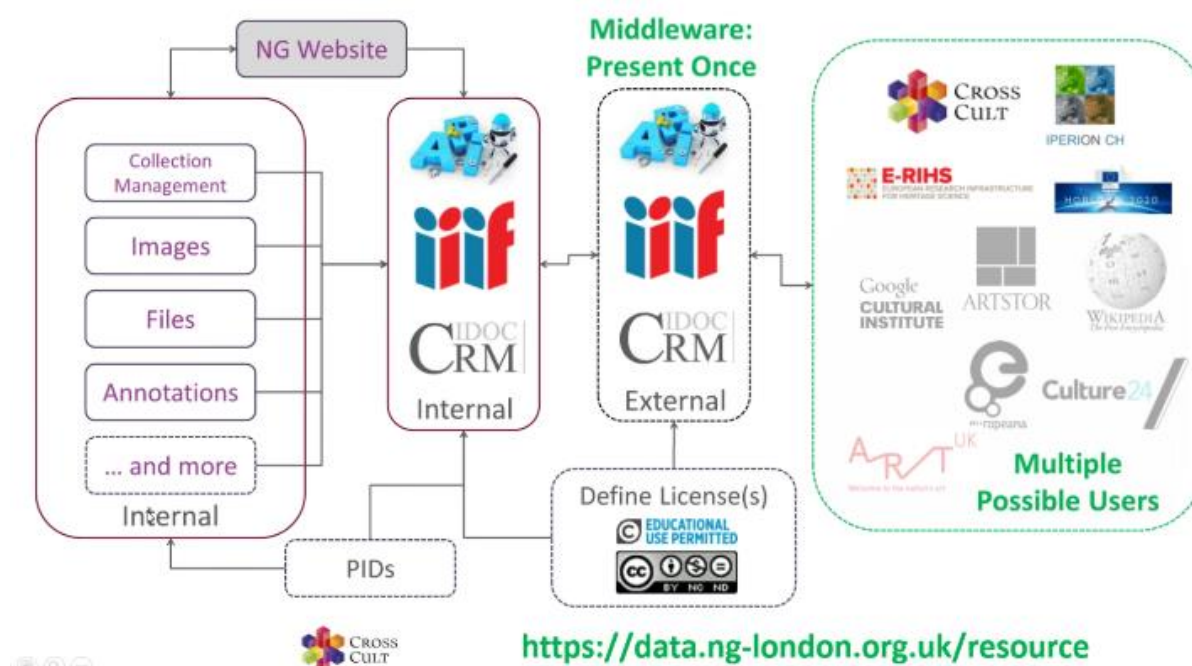
- Digital Object Identifier (DOI) doi:10.1186/2041-1480-3-9
- =handle + added features and rules for persistence, consistency, relationships, semantic interoperability....
- ISO standard, provided by registration agencies, such as <https://datacite.org/>

What will the benefits be from a future combination?

- Doi:10.XXXX/AAAA-AAAA-AAAA-AAAA

Фиг. 5. PID Системата на Националната галерия на Великобритания към юни 2020

National Gallery – Aggregating internal data for re-use



What's next? – Linked Open Data

- Explore how best to increase the awareness of NG published PIDs – **F**indable
- Implement a sustainable supportable middleware system, incorporating IIF and Persistent Identifiers (PIDs), with standard data end-points/APIs – **A**ccessible
- Integrate CIDOC-CRM and external open thesauri and vocabularies (eg Getty Vocabularies, WikiData) mappings and other external Linked Open Data resources – **I**nteroperable
- Assign and publish standard licences for all of our published data – **R**e-usable



Фиг. 6. LOD и агрегиране на вътрешните данни за повторна употреба –
НГ на Великобритания

В заключение от направеното детайлно проучване за същността, целите и прилагането на FAIR принципите за цифрови обекти представлящи културно съдържание може да се направи извода, че положителен сценарий в български условия, с наличните детайлни препоръки от ЕК, от практики и GLAM общността е възможен. Необходима е обаче много работа по изглаждане на детайлите и изясняване на целите по този сценарий в рамките на консорциума Клада-БГ. Доколко наличната технология и ясният план за управление на данните, който предстои да се създаде ще заработи в съгласие с доставчиците на културно съдържание и дали ще бъдат преодолені правно-административните пречки зависи от екипната воля за общи действия.

ИЗТОЧНИЦИ

FAIR ПРИНЦИПИ, FAIR УСЛУГИ, УПРАВЛЕНИЕ НА ДАННИ В ХРАНИЛИЩА И ИНФРАСТРУКТУРИ

- [1] **ОСНОВНИ ПРИНЦИПИ НА FAIR, 2016.**

<https://doi.org/10.1038/sdata.2016.18>; <https://www.go-fair.org/fair-principles/> Посл. посетен: 10-07-2020

- [2] Nielsen, Lars Holm. **ZENODO: FAIR DATA IN A GENERIC DATA REPOSITORY.**

[<https://zenodo.org/record/2650088#.Xt48wqZLepr>] Посл. посетен: 10-07-2020

- [3] **[SERVICES TO SUPPORT FAIR DATA](#), report and recommendations**

OpenAIRE, FAIRsFAIR, RDA Europe, FREYA, EOSC-hub. 2019 Посл. посетен: 10-07-2020

- [4] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al.

THE FAIR GUIDING PRINCIPLES FOR SCIENTIFIC DATA MANAGEMENT AND STEWARDSHIP.

Sci Data 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18> Посл. посетен: 10-07-2020

- [5] **GUIDELINES TO FAIRIFY DATA MANAGEMENT AND MAKE DATA REUSABLE –**

PARTHENOS project [http://www.parthenos-project.eu/portal/policies_guidelines] Посл. посетен: 10-07-2020

- [6] **DATA MANAGEMENT EXPERT [GUIDE](#)** [<https://doi.org/10.5281/zenodo.3820473>]

- ^[7] **TURNING FAIR INTO REALITY**, EU, 2018

[https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_0.pdf]

ISBN 978-92-79-96546-3 doi: 10.2777/1524 Посл. посетен: 10-07-2020

ПОСТОЯННИ ИДЕНТИФИКАТОРИ – ИНСТРУМЕНТ ЗА ОСИГУРЯВАНЕ НА FAIR УСЛУГИ

- [8] Dappert, A., Farquhar, A., Kotarski, R. and Hewlett, K., 2017.

CONNECTING THE PERSISTENT IDENTIFIER ECOSYSTEM:

BUILDING THE TECHNICAL AND HUMAN INFRASTRUCTURE FOR OPEN RESEARCH.

Data Science Journal, 16, p.28. DOI: <http://doi.org/10.5334/dsj-2017-028>

- [9] PID FOR HERITAGE COLLECTIONS [<https://zenodo.org/record/3630809#.Xt5nn9Uzapq>]PID:

<https://orcid.org/about/what-is-orcid/mission>

- [10]PID: <https://isni.org/> Посл. посетен: 10-07-2020

- [11]POWERING THE PID GRAPH: announcing the DataCite GraphQL API, May 6, 2020 by Martin Fenner,

<https://doi.org/10.5438/yfck-mv39> Посл. посетен: 10-07-2020

ПРИЛОЖЕНИЕ. IIIF – ИЗОБРАЖЕНИЯ В НАУЧНИ ИНФРАСТРУКТУРИ

Достъпът до ресурси, базирани на изображения, е от съществено значение за научните изследвания, науката и предаването на културни знания, особено в хуманитарните науки. Въпреки че дигиталните изображения съставляват голяма част от информационното съдържание в уеб достъпните портали от изображения, книги, вестници, ръкописи, карти, свитъци, колекции от един лист и архивни материали, те остават затворени в силози, като достъпът е ограничен с локални приложения. Причините за този ограничаващ изследванията факт са както правни неясноти, така и административни пречки, липса на национална политика. На портала на КЛАДА-БГ се предвиждат да се качат изображения от музеите и институциите –партньори с различен характер: част от публикация с необявен източник, част от сайта на доставчика, изображения от Интернет с КК лиценз, собствени изображения със защитено авторско право. Как всички те да могат да се използват некомерсиално, по унифициран начин да се преглеждат, да се анотират и пре-използват за целите на проекта? Възможно решение е ползването на IIIF **International Image Interoperability Framework** [<https://iiif.io>]. Към 2020 г. IIIF вече е регистрирана търговска марка, но възниква в резултат от усилията на разрастваща се общност на водещи световни изследователски библиотеки и хранилища за изображения за съвместно сътрудничество по създаване на оперативно съвместима технология и обща рамка за достъпване до дигитални образи¹¹.

Основните ползи от ползването ѝ са: Да даде на учените безпрецедентно ниво на еднакъв и богат достъп до ресурси, базирани на изображения, хоствани по целия свят; Да дефинира набор от общи интерфейси за програмиране на приложения, които поддържат оперативна съвместимост между хранилищата на изображения; Разработване, култивиране и документиране на споделени технологии, като сървъри за изображения и уеб клиенти, които предоставят потребителско изживяване от световна класа при гледане, сравнение, манипулиране и коментиране на изображения

IIIF ползи (английски):

- Advanced, interactive functionality for end users
- Fast, rich, zoom and pan delivery of images
- Manipulation of size, scale, region of interest, rotation, quality and format.

¹¹ Демо и приложения на IIIF [<https://iiif.io/apps-demos/>] Посл. посетен: 10-07-2020

- Annotation - IIIF has native compatibility with the W3C annotation working group's Web Annotation Data Model, which supports annotating content on the Web. Users can comment on, transcribe, and draw on image-based resources using the Web's inherent architecture.
- Assemble and use image-based resources from across the Web, regardless of source. Compare pages, build an exhibit, or view a virtual collection of items served from different sites.
- Cite and Share - IIIF APIs provide motivation for persistence, providing portable views of images and/or regions of images. Cite an image with confidence in stable image URIs, or share it for reference by others—or yourself in a different environment.
- Support for institutional authentication. IIIF is designed to support access control and can leverage existing SSO systems, via the IIIF Authentication API
- System flexibility. Use any IIIF-compatible software for viewing, comparing and manipulating images, regardless of the back-end server. Swap parts of the stack at any time, or run multiple components in parallel at once.
- Easy data transfer and sharing IIIF is standards-based, uses RESTful API construction, JSON-LD, and follows Web patterns, simplifying processes for data migration and sharing. Avoidance of vendor lock-in, complicated migrations, and system overhauls. IIIF provides the ability to separate image delivery user interfaces from back end data stores, allowing repositories the ability to update image servers and databases without changing front end delivery, or vice versa, avoiding the need to re-architect the entire stack. With an active and growing community of organizations developing and supporting IIIF-compatible technologies, there are plenty of software options to choose from. The work of implementing the IIIF APIs can be broken down into the core APIs (Image and Presentation) and additional APIs (Authentication and Content Search). The recommendation is to implement the core APIs first, and then the others can be added later depending on your needs. The Image API allows for interoperable delivery of digital images. Simply, the API allows for constructing parameterized URIs for getting a particular region and/or size of an image. This ability to get regions and sizes of the image is what supports delivering thumbnails as well as deep zoom viewing. The image API also specifies a way to get information (info.json) about the image and what image derivatives and image

manipulation features are supported for that image. What you need to implement is the server-side piece. За GUIDE с допълнително обяснение виж [ТУК](#).

- Reduction of long term total costs
- A rich ecosystem of interoperable software, including many high quality, open source options, keeps licensing and operational costs low and predictable over time.
- Combining content from across repositories
- Unify presentation of content from many different stores (within or outside your institution) without complex or expensive system or data migrations. Bring together content from your institutional DAM, library digital collections, and external sources.
- Publish once, re-use often. Deliver images from your own site for manifold uses; host a single copy and embed in other sites. No need to transfer images to others for them to locally load for one off analysis or republishing.
- A thriving global network. Join a global network of image suppliers making content available in a common , interoperable framework. Tap a growing suite of software tools and platforms. Maximize the use of your images on the Web. Unlock new potential with interoperability.