

Univerzita Karlova

Filozofická fakulta

Ústav informačních studií a knihovnictví

Informační studia a knihovnictví

Bakalářská práce

Martin Flohr

**Otevřený propojený katalog vědeckých přístrojů [Equipment.data.ac.uk](https://equipment.data.ac.uk)
Equipment.data.ac.uk – The Linked Open Catalogue of Scientific Equipment**

Poděkování:

Rád bych tímto poděkoval Dr. Janu Dvořákovi za odborné vedení mé bakalářské práce, za trpělivost, kterou měl s mými časovými (in)dispozicemi, za cenné rady, doporučení informačních zdrojů i informace přímo z akademického prostředí.

Prohlášení:

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

FLOHR, Martin. *Otevřený propojený katalog vědeckých přístrojů Equipment.data.ac.uk*. Praha, 2023. Bakalářská práce. Univerzita Karlova, Filozofická fakulta, Ústav informačních studií a knihovnictví. Vedoucí bakalářské práce Dr. Jan Dvořák.

Abstrakt: Cílem práce je obecný popis fenoménu sdílení vědeckého vybavení a konkrétně pak zmapování katalogu pro sdílení vědeckého vybavení Equipment.data.ac.uk a analýza jeho současného obsahu.

Dále je provedeno srovnání s projekty obdobného zaměření v ČR a zahraničí. Jsou zkoumány předpoklady a možnost realizace obdobného celostátního projektu v ČR.

Abstract: The aim of the work is a general description of the phenomenon of scientific equipment sharing and, more specifically, a mapping of the catalogue of shareable scientific equipment Equipment.data.ac.uk and an analysis of its current content.

Furthermore, a comparison is made with projects of a similar focus in the Czech Republic and abroad. The assumptions and the possibility of implementing a similar national project in the Czech Republic are examined.

Klíčová slova: sdílení vědeckého vybavení; výzkumné organizace; znalostní management; propojená data

Keywords: shared research equipment; research organizations; knowledge management; linked data

Obsah

1 Úvod.....	8
2 Sdílení vědeckého vybavení.....	10
2.1 Předpoklady.....	10
2.1.1 Vybavení.....	10
2.1.2 Prostor.....	11
2.1.3 Personál.....	11
2.1.4 Informační podpora sdílení.....	12
2.1.5 Financování.....	13
2.1.6 Podpora státu, legislativa, právní rámec.....	14
2.1.7 Motivace.....	15
2.1.8 Zájem o využití.....	15
2.2 Rozsah.....	16
2.3 Přínosy.....	17
2.4 Náklady a rizika.....	20
2.5 Technologie pro informační podporu sdílení.....	22
2.5.1 Katalog vybavení.....	23
2.5.2 Sběr dat.....	24
2.5.3 Otevřená propojená data.....	25
2.5.4 Zprostředkování transakce.....	28
2.5.5 Statistické výstupy.....	28
3 Případová studie: Equipment.data.ac.uk.....	30
3.1 Metodika.....	30
3.2 Metody sběru dat.....	30
3.3 O projektu Equipment.Data.....	32
3.4 Historie projektu.....	33
3.5 Procesy.....	34
3.5.1 Katalog.....	34
3.5.2 Předávání dat.....	35
3.5.3 Autodiscovery.....	36
3.5.4 Kontrolní výstupy.....	37
3.6 Datové formáty.....	37
3.6.1 Data na vstupu.....	38
3.6.2 Data na výstupu.....	40
3.7 Otevřená propojená data.....	41
3.8 Obsah katalogu Equipment.data.ac.uk.....	41
3.8.1 Databáze, počet záznamů a průběh v čase.....	41
3.8.2 Informace o vzorku dat.....	43
3.8.3 Výsledky analýzy dat.....	45
3.9 Silné a slabé stránky, perspektivy.....	47
4 Srovnání s obdobnými projekty.....	49
4.1 Obdobné projekty v zahraničí.....	49
4.2 Obdobné projekty v České republice.....	51
5 Diskuze.....	53
6 Závěr.....	58
7 Soupis použité literatury a dalších zdrojů.....	60

Definice pojmů a seznam použitých zkratk

Přístroj (equipment, instrument): Movitá věc technického rázu.

Zařízení (facility, large scale equipment): Budova, obsahující přístroje; případně přístroj takové velikosti, objemu nebo hmotnosti, že jej lze považovat za nemovitou věc, případně je samostatným obsahem budovy (např. jaderný reaktor, částicový urychlovač). Také ve významu instituce či její organizační složky.

Vybavení: Souhrnné označení pro přístroj nebo zařízení. Je to veškerý majetek technického rázu, který organizace používá v procesu výzkumu.

Utilizace, též vytížení, využití, míra využití... ve smyslu koeficientu využití: podíl doby, po kterou je vybavení aktivní, ku celkové době.

Core facility: Zařízení provozované akademickými institucemi, kde je koncentrováno různorodé vědecké vybavení určené pro využití jak v rámci instituce, tak pro externí subjekty. Toto zařízení není pořízováno ani vyhrazeno pro realizaci konkrétních výzkumů, ale právě pro účely sdíleného využití.

Skližení (harvesting): Proces, kdy jsou data z různých zdrojů (dostupných na internetu) pravidelně stahována pro účely dalšího zpracování nebo agregace do centrální databáze.

Autodiscovery: Proces, kdy si stroj sklízající data (harvester) automatizovaně zjišťuje dostupné zdroje dat.

OPD: Organisation Profile Document, soubor obsahující základní informace o instituci formou strukturovaných dat. Součástí tohoto souboru může být cesta k datům určeným ke skližení, v takovém případě OPD může sloužit pro realizaci autodiscovery.

Linked (open) data: Data volně přístupná na internetu ve strojově čitelném standardizovaném formátu která odkazují na další data (poskytující k datům kontext) a na která lze odkazovat pomocí HTTP URI.

1 Úvod

V rámci investic do vědy a výzkumu (ČR 2% státního rozpočtu, resp. 0,63% HDP v roce 2021^[1]) jsou významné částky investovány do nákupu vědeckého vybavení.

Je pochopitelným zájmem, aby takto pořízená technika byla využita maximálně efektivně. Jednou z možných strategií je sdílení vědeckého vybavení. Cílem je, aby byla maximalizována doba využití přístroje a minimalizována doba nečinnosti. Dále tato strategie omezuje duplicitní nákup přístrojů a nákup nepotřebných nebo málo využívaných přístrojů.

Pokud tedy organizace vlastní přístroj¹, který není využíván nepřetržitě, lze jej poskytnout k dalšímu využití, a to buď v rámci dalších projektů v rámci stejné organizace, nebo vně organizace, u dalších subjektů. To může v důsledku přinést nejen vyšší efektivitu, ale i úspory nepořízením dalšího přístroje (pokud volné kapacity stávajících přístrojů pokryjí poptávku po něm v rámci jiných projektů či organizací). Některá zařízení mohou cíleně vznikat přímo za účelem sdílení vědeckého vybavení.

Sdílení vědeckého vybavení přináší pochopitelně nejen výhody, ale i určitá rizika, kolizní body, které musí každá organizace před zapojením do sdílení vědeckého vybavení vyhodnotit.

Předkládaná práce se právě těmito aspekty zabývá.

V teoretické části, věnované sdílení vědeckého vybavení, jsou řešeny předpoklady pro realizaci, zhodnoceny přínosy i rizika této strategie, nastíněna otázka podpory ze strany státu a vůle a motivace organizací se na sdílení podílet. V závěru teoretické části je otevřena otázka organizačního zajištění, kterou následně rozvíjí praktická část práce.

¹ Pod pojmem "přístroj" si v tomto kontextu můžeme představit nejen určité technické vybavení, ale i celé objekty, jak je rozvedeno v následující kapitole.

Praktická část práce se zabývá formou případové studie konkrétním řešením pro podporu sdílení, otevřeným katalogem Equipment.data.ac.uk. Jsou popsány použité datové formáty, analyzován obsah dat projektu, posouzeny silné a slabé stránky projektu.

Následně jsou uvedeny obdobné projekty a je zhodnocena možnost implementace v ČR.

Cílem této práce je upozornit na fenomén sdílení vědeckého vybavení a přispět k diskusi o možnostech realizace v České Republice.

2 Sdílení vědeckého vybavení

Sdílení vědeckého vybavení je strategie, jak zvýšit jeho využití (utilizaci) tím, že jej vlastník poskytuje i dalším subjektům. Vybavení je tak využito i v době, kdy by jej vlastník nevyužíval a vybavení by zůstávalo nečinné.

2.1 Předpoklady

K uskutečnění myšlenky sdílení vědeckého vybavení je zapotřebí vytvořit vhodné podmínky.

2.1.1 Vybavení

Klíčovou komponentou sdílení je samotné vybavení. Zde je nutné odlišit dvě modelové situace: v prvním případě buduje výzkumná organizace (univerzita) nové pracoviště sdíleného vybavení (core facility) již s ohledem na sdílení vědeckého vybavení, ve druhém se výzkumná organizace zapojí do konsorciálního sdílení s již vybudovaným zázemím (které ne vždy musí respektovat výše uvedená doporučení) a nabídne své vybavení v rámci konsorcia.

V prvním případě je vhodné vzít již při plánování a pořizování vybavení v potaz potenciálně velkou a různorodou uživatelskou základnu. Zde je nutné dobře zvolit rovnováhu mezi uživatelsky přívětivým vybavením “denní potřeby”, které ocení většina uživatelů, a výhody špičkových přístrojů, které mohou inspirovat zcela nový a přelomový výzkum[2].

Nejde však jen o vybudování vhodného portfolia vědeckého vybavení. Je nutné průběžně zjišťovat potřeby uživatelské základny a reagovat na ně obnovováním, doplňováním a vyřazováním vybavení[2].

Dalším aspektem je forma sdílení z hlediska obsluhy přístroje. V některých případech jde o poskytnutí samotného přístroje, který si výzkumník (po potřebném proškolení) obsluhuje sám. Jinou variantou je provedení experimentu s asistencí podpůrného personálu či provedení

celého experimentu “na klíč”, formou nabízené služby. Zcela specifickým případem jsou pak přístroje, které lze obsluhovat na dálku a výsledek výzkumného procesu lze obdržet online. Takovým případem je například pokusný tokamak (komora pro termojadernou fúzi) Golem na FJFI ČVUT v Praze, kde je možné zadání parametrů výboje a výsledky jsou k dispozici online na adrese: <http://golem.fjfi.cvut.cz/shots/0/>. Obdobné systémy pracují i v zahraničí, jako například [3].

Zde sice nemůžeme hovořit o sdílení vědeckého vybavení (spíše o sdílení strojového času), ale i tímto postupem lze zvýšit utilizaci přístroje.

2.1.2 Prostor

Přestože prostor se může jevit jako zcela samozřejmý předpoklad, právě společný prostor, kde se setkávají odborníci různých profesí, může být významným benefitem sdílení vědeckého vybavení a podpořit výměnu znalostí napříč obory, jak uvádí například případová studie [4] zkoumající přínosy konceptu sdíleného vědeckého vybavení v oblasti biomedicíny a farmaceutického průmyslu. Nevhodně uspořádaný prostor může tento přínos oslabit nebo zcela i znehodnotit.

V případě tzv. core facilities by oproti tradičnímu členění prostoru dle potřeb jednotlivých výzkumných skupin měly být respektovány potřeby jednotlivých přístrojů [2]. Tento přístup by měl být zvážen i v případě prostor, které doposud nevyhovují sdílení vědeckého vybavení - optimálních podmínek pak může být dosaženo rekonstrukcí.

2.1.3 Personál

Klíčem k úspěšnému provozování core facility je technický personál, který má znalosti, je vstřícný a působí jako mentor pro uživatele. Vzhledem k tomu, že klienti platí za používání zařízení nebo za výsledky, lze na tato zařízení pohlížet spíše jako na firmy a důležitá je zde tedy interakce s klientem [5]. Zde je patrný rozdíl oproti personálu běžné laboratoře. Edukační roli personálu core facility zmiňuje i [2], když uvádí, že zaměstnanci působí jako institucionální kurátoři znalostí v oblasti analytických metod a schopností daných přístrojů.

Vysoké nároky na personál jsou podle autorů dány tím, že škála uživatelů core facility velmi široká. Proto musí být zaměstnanci zařízení zapojeni do školení uživatelů přizpůsobiví ve svých metodách.

Otázka personálu se dotýká i zajímavé problematiky spoluautorství podpůrného personálu na realizovaných a publikovaných výzkumných výsledcích. Instituce by měla prosazovat vytváření a zavádění kodexů chování pro uživatele, které by nabádaly k uznávání technického personálu v publikacích, samozřejmě s ohledem na pokyny pro autory (časopisu nebo příslušné odborné společnosti). Tyto pokyny obvykle respektují jejich odborný příspěvek k výzkumu a zaručují spoluautorství, současně také požadují, aby se rezidentní expert zodpovídal technické komunitě. Pokud není spoluautorství vhodné, pak by mělo být normou uvedení zaměstnance jménem, nikoliv jen názvem zařízení jako celku, za předpokladu, že jejich příspěvek přesahuje základní školení a rutinní laboratorní dozor[2].

2.1.4 Informační podpora sdílení

Pro efektivní sdílení vědeckého vybavení nestačí jen samotná existence vybavení a vůle jej sdílet. Nezbytné je také, aby potenciální uživatelé měli možnost se dozvědět o dostupných příležitostech a službách. Toto je výzvou zejména vně instituce, i když problém s nedostatečnou komunikací a z toho plynoucí nízkou utilizací vybavení může nastat i uvnitř instituce. Za tímto účelem je potřeba vybudovat nějakou formu katalogu sdíleného vybavení.

Za klíčovou výzvu považuje znalost dostupných sdílených služeb a technologií v rámci core facility i ABRF². Zjišťování informací o core facility může být podle nich mimořádně obtížné a je potřeba, aby instituce provozující zařízení typu core facility zhodnotily svá portfolia a zajistily vývoj katalogů a jejich propagaci jak uvnitř, tak vně své instituce[6].

Na propagaci sdíleného vybavení a služeb navazuje podpora procesů spojených se sdílením:

² Association of Biomolecular Resource Facilities, <https://abrf.org/>. Asociace biomolekulárních výzkumných zařízení je mezinárodní organizace sdružující 340 laboratoří typu core facility z 41 zemí.

- rezervace
- objednávky služeb
- platby
- dokumentace proběhlých využití a jejich případná publikace, plně či selektivně

Zatímco výše zmíněná propagace je nezbytností, podpora návazných procesů je doplňkem, který může šetřit administrativu.

Pro oba uvedené účely lze zvolit odlišná řešení od primitivních (webová stránka či tabulka ke stažení a rezervace e-mailem) až po sofistikovaná softwarová řešení.

Tomuto tématu je věnován ještě samostatný oddíl 2.5.

2.1.5 Financování

Naprostým základním předpokladem je financování. V případě sdíleného vybavení mohou nastat dvě modelové situace: sdílení již existujícího vybavení nebo vybudování nového pracoviště se sdíleným vybavením, např. tzv. core facility. V prvním případě jde tedy zejména o financování provozních nákladů, v druhém ještě navíc o rozsáhlou investiční akci.

Bez řádného financování provozních nákladů (spotřební materiál, média, energie, náklady na prostory, amortizace, pojištění) a personálních nákladů (technický personál, administrativa) by chyběla motivace ke sdílení vědeckého vybavení.

Tyto náklady jsou alespoň zčásti hrazeny příspěvkem uživatele (v případě individuálních výzkumníků je obvykle hrazen z grantu jako součást nákladů na realizaci výzkumu), další část může být dotována domovskou institucí, v některých případech z (institucionálních) grantů[2].

V případě budování nové core facility je investiční akce financována ze státního rozpočtu (např. formou účelové podpory, státního programu), nestátních finančních prostředků (darů,

fondů), případně kombinací těchto možností.

V případě, že se instituce nebo konsorcium uchází o prostředky ze státního rozpočtu, může být fakt, že se jedná o sdílené zařízení, benefitem zvyšujícím šanci na přidělení prostředků, protože pravidla těchto řízení zpravidla požadují demonstraci efektivních zásad řízení a rozsáhlou uživatelskou základnu, což u dobře vedeného pracoviště se sdíleným vybavením vyplývá z jeho principu[2].

2.1.6 Podpora státu, legislativa, právní rámec

Sdílení vědeckého vybavení je jednoznačně v zájmu státu. Sdílením vědeckého vybavení je podpořeno efektivní využití prostředků vynaložených ze státního rozpočtu do nákupu vědeckého vybavení. Dalším aspektem je, že sdílení vědeckého vybavení má pozitivní vliv na mezioborovou spolupráci, přípravu mladých vědců na další kariéru a rozvíjí jejich schopnosti v práci s přístrojovým vybavením. Celkově lze tedy úsilí vložené ze strany státu do podpory sdílení vědeckého vybavení považovat za podporu vědy a výzkumu. Podrobněji je tato problematika zpracována v oddíle 2.3 Přínosy.

Stát by měl tedy vytvářet příznivé podmínky pro sdílení vědeckého vybavení, tedy legislativu, která bude tyto aktivity podporovat (nebo alespoň umožňovat, minimálně nebránit). Zásadní je legislativa související s financováním vědy a výzkumu.

Jako příklad lze uvést aktivity NIH³, který aktivně podporuje sdílené využívání zařízení typu core facility a stanovil zásady, které umožňují sdílení. Zásady jsou zveřejněny na webu NIH v sekci grantů a financování[6].

Podporu státu lze demonstrovat i na nizozemském státem podporovaném fondu Mibiton, který má umožnit přístup k vědeckému vybavení pro začínající R&D společnosti, zejména

³ The National Institutes of Health, národní institut zdraví, součást amerického ministerstva zdravotnictví a sociálních věcí.

biotechnologické firmy. Program “Share” tohoto fondu cílí na spolupráci startupů, inovativních malých a středních podniků a výzkumných institucí za účelem sdílení vědeckého vybavení[7].

Další právní rámec pak zajišťují smluvní dokumenty. V případě core facility půjde zejména o pravidla pro fungování takového pracoviště, tedy vztahy provozovatel - uživatel (práva, povinnosti, zodpovědnost, úhrada nákladů). V případě konsorciálního je právní rámec zajištěn formou smluvních dokumentů (dohoda členů konsorcia).

2.1.7 Motivace

Prvním předpokladem k realizaci sdílení vědeckého vybavení je motivace vlastníka ke sdílení. Pokud pomineme čistý entuziasmus, motivací pro sdílení je buď přímý ekonomický přínos (příjem ze sdílení převyší náklady na sdílené využití) či reciproční služba (připojením do konsorciální sítě účastník získává i možnost využívat přístroje dalších účastníků), nebo přínos v oblasti vědy (účast na určitém výzkumném projektu, podíl na určitém vědeckém výsledku).

2.1.8 Zájem o využití

Na druhé straně, tedy u potenciálních uživatelů sdílení musí existovat zájem o využití vybavení, tedy poptávka. V případě, že bude vlastník sdílet zařízení, které nemá pro vědeckou komunitu význam, nebude účel sdílení splněn, neboť nedojde ke zvýšení utilizace. Zájem o využití lze do jisté míry ovlivnit jak pozitivně, tak negativně. Pozitivně lze působit například propagací nabídky sdílení (ať již aktivně, nebo jen jejím zařazením do relevantního katalogu), přizpůsobením přístrojového portfolia skutečným potřebám cílové skupiny (např. realizací průzkumu potřeb), ale i vytvořením příznivého prostředí v rámci daného zařízení, a to jak po stránce materiální, tak personální. Nedocenitelná je v této oblasti osobní zkušenost výzkumníků předávaná kolegům[2].

Podrobně se otázkou spokojenosti uživatelů zabývá studie [8]. Autoři studie (Gyeonggi Bio-Center) stanovili požadavky na kvalitu služby v těchto základních kategoriích:

- spolehlivost

- důvěra
- materiální záležitosti
- empatie (personálu)
- schopnost reagovat na požadavky

Negativní vliv na zájem o využití sdílení může mít mimo absence výše zmíněných prvků také špatné nastavení pravidel pro využívání sdíleného vybavení nebo nepříznivá cenová politika.

2.2 Rozsah

Sdílení vědeckého vybavení lze realizovat v rámci instituce v různém rozsahu. Zařízení může vzniknout přímo za účelem sdílení, dále může být sdíleno vybavení již existujících zařízení pro zvýšení jejich utilizace, případně může být sdílení jen marginálním jevem v rámci aktivit dané instituce.

- A) **Core facility.** Sdílení vědeckého vybavení v rámci jednoho pracoviště (často univerzity), případně se jedná o jedno společné pracoviště několika organizací (např. BIOCEV jako společný projekt AV ČR + UK). Core facility vzniká přímo za účelem sdílení vědeckého vybavení, je to jeho primární účel. Slouží jak výzkumným pracovníkům instituce, tak externím uživatelům. Některá poskytují školení pro práci s instalovanou technikou, většina také nabízí výpočetní a statistické služby, vše pod jednou střechou.
- B) **Konsorciální** sdílení vědeckého vybavení předpokládá vytvoření konsorcia výzkumných organizací (nebo i soukromých subjektů), které se smluvně dohodnou na vzájemném sdílení vědeckého vybavení. Konsorciální sdílení vyžaduje vždy nějaký řídicí subjekt, a to buď jednu (zpravidla nejvýznamnější) organizaci v rámci konsorcia, nebo externí organizaci, kterou konsorcium pověří administrací sdílení. Sdílení se zde odehrává jako cesta k vyššímu využití stávajících přístrojů, ale samozřejmě přináší také benefity na poli vědecké spolupráce, stejně jako u core

facility.

Specifickým druhem konsorciálního sdílení je pak sdílení **celostátní**, kdy roli administrátora převezme stát, jeho organizační složka nebo subjekt státem pověřený. Stát určí legislativou podmínky fungování takové celostátní sítě. Účast v takové síti může být dobrovolná, nebo povinná. Podobné sítě vznikají i v **celoevropském nebo celosvětovém měřítku**.

- C) Ad hoc sdílení přístroje je případ, kdy sdílení není výhradním, ani hlavním záměrem, ale spíše přidruženým jevem. Instituce používá přístroj a v určitém okamžiku vznikne poptávka po sdílení, případně je shledána příležitost ke sdílení, která je následně realizována.

Příkladem informační podpory formou katalogu pro konsorciální sdílení v celostátním rozsahu ve Velké Británii je projekt Equipment.data.ac.uk, kterým se zabývá praktická část této práce.

2.3 Přínosy

Základní myšlenkou sdílení vědeckého vybavení je úspora nákladů: S rostoucím využitím (utilizací⁴) přístrojů klesá jejich nutný počet pro splnění stejného objemu práce. Z ekonomického pohledu roste i efektivita vynaložených prostředků: ze stejného objemu vynaložených prostředků získáváme větší užitek.

Jak uvádí [4]: Sdílené zdroje se obvykle používají, když nejsou k dispozici prostředky pro nákup přístroje pro každou laboratoř. Jediný přístroj pak bývá umístěn a zpřístupněn výzkumným pracovníkům napříč odděleními nebo celou organizací.

4 Utilizaci definujeme jako procentuální poměr času využití přístroje k součtu času využití přístroje a času zahálení přístroje. V případě 100% utilizace je přístroj využíván nepřetržitě.

Z opačného pohledu je pak možné se stejným rozpočtem pořídit namísto většího počtu identických přístrojů širší škálu různorodých přístrojů, což ocení uživatelská základna core facility nebo konsorciálního sdílení. V neposlední řadě se tak otevírá přístup k nákladnému špičkovému vybavení i uživatelům (individuálním výzkumníkům z řad studentů i pro malé a střední podniky zabývající se výzkumem či vývojem), kteří by si jej jinak nemohli dovolit[2]. Autoři dále uvádí, že jde o přínos na obou stranách: univerzity těží z existence core facility již jen tím, že jejich vybavení je dobře dostupné, správně rozvržené a viditelné potenciálním uživatelům. Přínosy vidí i v oblasti pořizování vybavení: eliminují se duplicitní nákupy, využití zakoupeného vybavení je maximalizováno. Správa krátkodobých i dlouhodobých výdajů je snazší se širokým portfoliem vybavení a personálu. Přínosy shledává i v oblasti propagace vybavení díky širší uživatelské základně.

Další přínosy jsou konstatovány v oblasti spolupráce s dodavateli. Dodavatelé navazují s provozovateli sdíleného vybavení dlouhodobé vztahy, z čehož plyne nejen příznivý dopad na ceny produktů a údržby, ale také na rozšířené služby. Provozovatel core facility je pro dodavatele také perspektivním partnerem z hlediska propagace: rozsáhlé a různorodé publikum uživatelů je vhodným cílem workshopů a krátkých kurzů[2].

Mimo těchto relativně přímých přínosů lze však konstatovat mnohé nepřímé přínosy: přítomnost mnoha vědeckých přístrojů z různých oborů ve společném prostoru napomáhá setkávání vědců z mnoha oborů či týmů. V tomto smyslu může být koncept sdíleného vybavení součástí designu, který podporuje kontakt, spolupráci a kreativitu[4].

Podrobněji se tato studie věnuje právě přínosu sdílení vědeckého vybavení ke zlepšení mezioborové spolupráce, výměně znalostí a zlepšení interní komunikace, a to prostřednictvím tří případových studií z firem působících v oblasti biomedicíny a farmacie. Závěrem studie autoři uvádějí, že sdílené vybavení může mít znatelné přínosy v této oblasti, protože podporuje komunikaci umístěním vědců z různých týmů, oborů a větví organizační hierarchie do častého vzájemného kontaktu[4].

Core facility slouží jako bod pro společnou komunikaci mezi výzkumnými skupinami a

přitahuje spolupráci s průmyslem i pracovní příležitosti. Univerzity pak napomáhá svou atraktivitou získávat nové postgraduální studenty i postdoktorandy, stejně jako nové akademické pracovníky. Výzkumníci na oplátku získávají znalosti v analytických metodách a mají k dispozici kvalifikovaný personál, servis a školení[2].

Vlivem prostorové koncentrace a sdíleného využití vědeckých přístrojů na výkonnost výzkumu se zabývá i studie [9]. Autoři dospěli k závěru, že sdílené vědecké vybavení poskytuje více příležitostí pro úspěch ve výzkumu, pomáhá mladým výzkumníkům bez výzkumného vybavení a také může přispět k zvýšení utilizace a brání případům, kdy jsou přístroje málo využité nebo jsou zcela nevyužité. Dále autoři uvádí, že prostřednictvím sdílení výzkumného vybavení lze vytvořit ekosystém a platformu za účelem propojení nabídky a poptávky služeb. To může přispět k tvorbě výzkumných výsledků bez nutnosti pořizovat výzkumné vybavení. Jako další přínos zmiňují již dříve uvedený fakt, že sdílením vědeckého vybavení se snižují potřebné výdaje a zvyšuje se účinnost výzkumu. Vyzdvihují také mezioborovou spolupráci výzkumných organizací a její příznivé dopady, např. vytváření sítí výzkumných pracovníků, což může globálně přispívat k podpoře výzkumné činnosti, sdílení výsledků výzkumu a novým objevům. Autoři tak v zásadě shrnují závěry předchozích studií a potvrzují uvedené přínosy sdílení vědeckého vybavení.

Pokud tedy shrneme doposud uvedená zjištění, sdílení vědeckého vybavení přináší:

A) přínosy ekonomické

- vyšší utilizaci, tedy efektivitu vynaložených prostředků
- úsporu z principu pořízení menšího počtu zařízení nebo možnost za stejné prostředky pořídit širší škálu přístrojů
- dosažení lepších cen a výhodnějších podmínek u dodavatelů

B) přínosy ve veřejném zájmu, ve vědecké rovině

- zpřístupnění špičkového vědeckého vybavení i těm, kteří by jinak neměli možnost jej využít

- díky prostorové koncentraci zlepšení mezioborové spolupráce, předávání znalostí, což prokazatelně zvyšuje efektivitu výzkumné činnosti
- dostupnost technického personálu (v případě core utility) přispívá k edukaci uživatelů ohledně přístrojového vybavení, případně jim přímo zajišťuje podporu
- uživatelé působící v rámci core facility jsou díky kontaktu s víceoborovým prostředím a díky dostupnosti technického personálu lépe připraveni na další výzkumné aktivity i na budoucí profesní uplatnění

2.4 Náklady a rizika

Sdílení vědeckého vybavení nepřináší jen výhody. Přináší zvýšené náklady a jsou s ním spojená rizika.

Využití zařízení externím subjektem představuje zvýšené náklady na energie, spotřební materiál, média, zvyšuje personální náklady na odborný personál. Také administrace sdílení představuje další personální i materiální zátěž - například opakující zadávání vybavení do katalogu[10].

Tyto náklady by měly být kompenzovány určitým příspěvkem ze strany uživatele, který je však nemusí pokrývat zcela[2]. Nelze tedy vyloučit, že sdílení bude i po odečtu příspěvků představovat pro instituci finanční zátěž. Tu by měly vyvážit nemateriální přínosy sdílení.

Vybavení, které bude sdíleno, je vystaveno větší zátěži (v důsledku vyšší utilizace). Vliv vyšší zátěže na životnost vybavení je nutné zvážit v koncepci sdílení. Na tento jev se lze ovšem dívat i tak, že díky kratší životnosti bude možná častější obměna zařízení.

Dalším negativním faktorem může být výskyt cizích pracovníků v prostorách organizace, pokud což sice přináší výše uvedené přínosy ve sdílení znalostí, ale také představuje určité bezpečnostní riziko. Toto se netýká případů sdílení bez osobní přítomnosti (online, formou služby).

Přestože sdílení znalostí jako jeden z efektů sdílení vědeckého vybavení je pozitivní jev, za určitých okolností může vznikat mezi uživateli sdíleného vybavení obava z konkurence či obava o zachování určitého know-how[11][12].

V rámci sdílení vědeckého vybavení může docházet nejen k opotřebení, ale i poškození či zničení vybavení. S touto eventualitou musí počítat smluvní podmínky (pravidla), tak aby byly ochráněny investované prostředky[11], na druhou stranu nelze přikládat k tíži uživatele náklady na opravu v případě, že k závadě došlo v důsledku běžného opotřebení.

S tím souvisí i potřeba řešení situace, kdy se v důsledku závady zařízení stane nedostupné i pro další uživatele či vlastníka a z toho plynoucích důsledků (nutnost nalézt záložní řešení pro plánovaný výzkum, v extrémním případě i nemožnost jeho provedení)[11].

V případě sdílení konsorciálního typu nebo ad hoc sdílení, kdy zařízení mimo sdílení slouží i vnitřním potřebám instituce, lze jako negativum zmínit fakt, že v určitou dobu zařízení nebude k dispozici, což je potřeba řešit plánováním a dobrou organizací práce. V každém případě takové zařízení nebude možné nasazovat operativně.

Pro uživatele sdíleného vybavení se může jevit jako negativní faktor také určitá nejistota ohledně priorit, kdy nelze předpokládat, že určitý tým nebo výzkumník bude mít přednostní přístup k přístroji (jako by měl v případě, že by přístroj byl pořízen cíleně pro potřeby konkrétního výzkumu)[2].

Problémem technického rázu pak může být v některých oborech uchování mezistavů výzkumu (např. vzorků) po dobu, kdy přístroj využívá jiný tým.

2.5 Technologie pro informační podporu sdílení

Jak již bylo uvedeno v oddílu [Informační podpora sdílení](#), pro efektivní sdílení vědeckého

vybavení je nutné, aby se potenciální uživatelé dozvěděli o dostupných zařízeních a službách. Dále je žádoucí existence technického řešení pro administraci sdílení, tedy ze strany uživatele rezervaci termínů, kontrolu dostupnosti a ze strany vlastníka plánování kapacit. Přispívá to k celkové spokojenosti uživatele.

Tyto požadavky může splnit i poměrně primitivní řešení, kdy je katalog prezentován tabulkou na webu či ke stažení a administrace požadavků i plánování využití se odehrává prostřednictvím e-mailu či sociálních sítí.

Níže se budeme zabývat pokročilejšími možnostmi, kdy funkci katalogu i administrace plní informační systémy.

V současnosti jsou dostupná jak komerční tak open-source řešení pro správu sdíleného vybavení, zahrnující rezervace přístrojů či termínů školení, účtování nákladů a rozúčtování plateb. Tyto nástroje pomáhají redukovat administrativu, mohou zjednodušit proces rezervace (což vede ke spokojenějším uživatelům) a především zvýšit efektivitu celého procesu. Příjmy z vyšší efektivity pak mohou vynahradit výdaj na investici do příslušného informačního systému[2].

Instituce mohou tyto funkce také integrovat do svých stávajících informačních systémů (zejména jako součást CRIS⁵ / RIMS⁶ řešení). Příkladem takového řešení mohou být například komerční produkty PURE společnosti Elsevier, Converis společnosti Clarivate, nebo Symplectic Elements společnosti Digital Science. Z open-source produktů lze zmínit DSPACE-CRIS.⁷

Specifická situace nastává v případě konsorciálních sdílení. Každý z účastníků konsorcia má

5 Current Research Information System

6 Research Information Management System

7 Výše uvedené produkty CRIS / RIMS jsou koncipovány pro integraci dat z mnoha zdrojů a inzerují širokou konfigurovatelnost. Skutečné nasazení těchto produktů pro podporu sdílení vědeckého vybavení se však podařilo potvrdit pouze u produktu Pure společnosti Elsevier, a to na University of Brighton.

své vlastní informační systémy. V případě katalogu vybavení je tedy potřeba najít řešení, jak data z různých informačních systémů (pro evidenci majetku) efektivně sdílet.

2.5.1 Katalog vybavení

Katalog vybavení je místem, kde se uživatel setkává s celým ekosystémem sdíleného vybavení. Dojem z katalogu tedy může významně ovlivnit celkové vnímání služby. Jeho vývoj by tedy (stejně jako vývoj libovolného uživatelského rozhraní) měl stát na dobré znalosti uživatelských potřeb. Níže lze tedy uvést jen některé obecné předpoklady, konkrétní katalog bude vždy přizpůsoben oborovým či lokálním specifikům.

Katalog musí poskytovat zejména možnost vyhledávat vybavení podle vhodných kritérií - ty musí vycházet z údajů obsažených v databázi vybavení. Dále lze poskytnout procházení hierarchickou strukturou vybavení, pokud je tato v databázi vybavení obsažena.

Ze základních kritérií pro vyhledávání lze uvést:

- název přístroje
- model přístroje
- výrobce přístroje
- plnotextové vyhledávání ve všech dostupných údajích (zahrne tedy i popis)

Výsledky vyhledávání by mělo být možné dále filtrovat podle kritérií, která dávají smysl s ohledem na strukturu poskytovaného vybavení. Cílem je redukovat výslednou sadu výsledků vyhledávání.

Ze základních kritérií pro filtrování:

- obor / kategorie přístroje (vyžaduje určitou klasifikaci, ideálně na základě standardizovaných taxonomií⁸ či heslářů⁹)

⁸ Např. konsorcium N8 používá ve svém katalogu <https://www.n8equipment.org.uk/> jednoduchou tříúrovňovou hierarchii vybavení (Class → Order → Genus, tedy Třída → Řád → Druh). Navzdory pojmenování úrovní není katalog omezen jen na biologické obory.

⁹ Např. CPV kódy (Common Procurement Vocabulary), zavedené Evropskou Unií nařízením Evropského parlamentu 2195/2002, <https://eur-lex.europa.eu/legal-content/CS/TXT/PDF/?uri=CELEX:32002R2195>) pro

- typ obsluhy (samostatně / s asistencí / formou služby / vzdáleně)
- dostupnost (je/není aktuálně/v zadaném termínu dostupný - pokud je tato informace v systému k dispozici)
- konkrétní instituce a nižší organizační jednotky (v případě konsorciálního sdílení)

Kritéria pro vyhledávání a filtrování mohou splývat, např. v konsorciu N8¹⁰ je klasifikace přístroje a konkrétní instituce součástí kritérií pro vyhledávání.

K vyhledané jednotce vybavení by měl katalog zobrazit minimálně tytéž informace, jaké jsou použity jako kritéria pro vyhledávání a filtrování vyhledaných výsledků.

Dále je vhodné zobrazit bližší specifikaci přístroje:

- podrobnější popis (text)
- technické parametry (ideálně formou strukturovaných dat, sada parametrů relevantní k danému oboru a typu přístroje, v praxi však bude s ohledem na snížení nároků na zadávání dat ze strany poskytovatele nejspíše nahrazena právě jen podrobnějším popisem)
- informace o možnosti sdílení (lokalitu, podmínky, kontakt, případně rovnou možnost zadat požadavek)
- fotografie přístroje

Katalog je realizován jako www služba, kdy webový server pracuje s daty databázového serveru. V zásadě rozlišujeme dvě rozhraní: frontend a backend.

Frontend: Rozhraní pro uživatele, poskytující vstup pro vyhledávání a výstup vyhledávání (seznam položek vyhovujících zadání).

Backend: Rozhraní pro poskytovatele sdílení (zadávání dat a administrace).

předměty veřejných zakázek.

10 Konsorcium N8 (osm významných univerzit ze severu Velké Británie), <https://www.n8equipment.org.uk/>

2.5.2 Sběr dat

Plnění katalogu by měl být proces, který bude zúčastněné organizace co nejméně zatěžovat.

V případě, že není inventář vědeckého vybavení evidován v žádném stávajícím IS, je nutné ruční zadávání vědeckého vybavení do databáze katalogu (v backend rozhraní), což představuje nepopulární administrativní zátěž[10].

Pokud je majetek organizace včetně vědeckého vybavení evidován v nějakém IS (evidence majetku, CRIS), je vhodným řešením jeho zpřístupnění / publikování exportem v předem dohodnuté struktuře a formátu. Tím dojde k zjednodušení procesu a úspoře agendy.

V případě konsorciálního sdílení pak tato data musí být nějakým způsobem doručena do centrálního úložiště. Možné varianty jsou např.:

- bez automatizace - zaslání pracovníkem organizace agregátorovi např. e-mailem
- nahrání souborů (csv, xml, tabulkové procesory) na vyhrazený (s)FTP server na straně agregátora nebo členské organizace
- pravidelné sklizení (harvesting) souborů (csv, xml, tabulkové procesory) uložených na předem dohodnutém místě v rámci struktury webu členských organizací[13]
- předání dat mezi IS v dohodnutém formátu API funkcí (náročné na konstrukci IS na straně serveru i na exportní funkce na straně IS u členů konsorcia)

Harvesting přináší výhodu v minimální náročnosti (vyloučení lidského faktoru) na straně člena konsorcia - data jsou automaticky stahována v pravidelných intervalech. Přitom není nutné budovat na straně člena konsorcia žádný sofistikovaný systém[14]. Obdobně lze automatizovat i případný upload dat na vyhrazený (s)FTP server.

Nutná je jen počáteční registrace člena konsorcia, kde jsou předány souhrnné informace o instituci a zejména lokace, ze které budou data pravidelně stahována.

2.5.3 Otevřená propojená data

Tim Berners-Lee, tvůrce world wide webu a propagátor propojených dat, navrhl pětihvězdičkovou škálu otevřenosti dat:

- jedna hvězdička za zpřístupnění dat na webu
- dvě hvězdičky, pokud jsou data ve strojově čitelném formátu
- tři hvězdičky, pokud jsou data v otevřeném (neproprietárním) formátu
- čtyři hvězdičky, pokud lze na data odkazovat (mají URI)
- pět hvězdiček, pokud je uveden i kontext¹¹, a to propojením na jiná data

V případě, že data získají všech pět hvězdiček, hovoříme o otevřených propojených datech[15].

Principy propojených dat:[16, s. 11]

- URI¹² pro pojmenování věcí.
- HTTP¹³ URI, aby bylo možné na URI odkazovat.
- Při volání URI by měl být vrácen výsledek
 - ve standardizovaném formátu (RDF, SPARQL)
 - výsledek by měl obsahovat užitečný obsah
- Obsažená data odkazují na další data.

Principem propojených dat tedy je, že zdroj dat má přiděleno HTTP URI, aby na něj bylo možné odkazovat (přistupovat k němu na internetu), na základě volání HTTP URI vrací výsledek ve standardizovaném strojově čitelném formátu, obsahuje relevantní data (vzhledem k URI) a součástí dat jsou odkazy na další data, zejména ve smyslu kontextu dat.

¹¹ Kontext: informace o tom, jaký význam data mají, tedy o jakou jde veličinu, jakou má jednotku; pokud jde o kód, z jakého je číselníku. Například hodnota 37 může být cokoliv. Teprve s doplněním kontextu (teplota, ve stupních celsia) získává publikovaný údaj význam. Obdobně SPZ vozidla, PSČ, nebo ADR kód.

¹² URI, Uniform Resource Identifier, definován standardem RFC 3986 Internet Engineering Task Force (IETF), <https://www.ietf.org/>. Standard dostupný z: <https://datatracker.ietf.org/doc/html/rfc3986>.

¹³ HTTP, Hypertext Transfer Protocol, definován standardem RFC 2616 Internet Engineering Task Force (IETF), <https://www.ietf.org/>. Standard dostupný z: <https://datatracker.ietf.org/doc/html/rfc2616>

Myšlenka propojených dat je realizována pomocí datového modelu RDF¹⁴. RDF je metadatový model, používaný pro popis dokumentů (informačních zdrojů). Je zamýšlen jak pro strojové zpracování, tak pro lidské čtení.

RDF reprezentuje dvojici objektů a vztah mezi nimi jako trojici entity - attribute - value (entita - vlastnost - hodnota), ačkoliv pro potřeby propojených dat hovoříme spíše o trojici subject - predicate - object (subjekt – predikát – objekt nebo podmět – vlastnost – předmět) [\[16, s. 17\]](#).

RDF tedy vylepšuje propojenost webu (hypertext) o použití URI k pojmenování vztahu mezi věcmi a také obou konců odkazu (výše uvedené trojice). Pomocí tohoto jednoduchého modelu umožňuje strukturovaná a polostrukturovaná data kombinovat, zpřístupňovat a sdílet napříč různými aplikacemi [\[17\]](#).

Právě tato vlastnost RDF jej činí ideálním pro aplikaci propojených dat.

Podstatné je, že atribut (predikát, vlastnost) není volně tvořený, ale pochází z definovaných slovníků či ontologií, která jsou označována jako schéma (též jmenný prostor). Ze známých schémat lze uvést například Dublin Core¹⁵, MODS¹⁶ nebo schema.org¹⁷. Protože i atribut má přiděleno URI, je zajištěn i odkaz na definici vztahu a další informace poskytující kontext [\[16, s. 38\]](#).

RDF je pouze abstraktní model. Zhmotnění RDF do konkrétního souboru (syntaxe) má několik variant:

¹⁴ Resource Description Framework, rodina specifikací vytvořených World Wide Web Consortium (W3c), <https://www.w3.org/>

¹⁵ Dublin Core, <https://www.dublincore.org/> je metadatové schéma vytvořené pro popis webových objektů, používané i pro popis knih, obrazů a hudby.

¹⁶ MODS, Metadata Object Description Schema, <https://www.loc.gov/standards/mods/> je metadatové schéma pro popis bibliografických elementů, vytvářené Library of Congress Spojených států amerických.

¹⁷ Schema.org, <https://schema.org/> je aktivita založená společnostmi Google, Microsoft, Yahoo and Yandex, nyní komunitního charakteru, s cílem vytvářet, udržovat a propagovat schémata pro strukturovaná data na internetu.

- Turtle (ttl) je standard W3C pro zápis RDF, snadno lidsky čitelný i zapisovatelný.
- RDF/XML je zápis RDF ve formátu XML¹⁸.
- RDFa¹⁹ umožňuje realizaci propojených dat přímo vkládáním RDF do kódu HTML5.

Příkladem využití propojených dat při sdílení vědeckého vybavení může být OPD²⁰ jako zdroj souhrnných dat o instituci, mimo jiné i odkazu na data o vědeckém vybavení pro harvesting. Odkaz na RDF soubor obsahující OPD je vložený v html hlavičce domovské stránky instituce[14].

OPD tak plní definici otevřených propojených dat, jde o na webu zpřístupněná, strojově čitelná data v otevřeném formátu, na OPD lze odkazovat a OPD obsahuje informace o kontextu formou odkazů.

Praktickým přínosem takového řešení je, že v případě změn v údajích o organizaci nebo umístění souboru s daty pak postačí změna v OPD, systém si načte změny v nejbližší sklizni. Není nutné nic hlásit provozovateli systému. Tento proces automatického zjišťování zdroje dat je označován jako data autodiscovery[13][14].

V tomto případě tak k výše uvedené škále otevřených dat OPD přidává snadnou “objevitelnost” (discoverability), tedy schopnost dat být automaticky strojově nalezena (a následně sklizena).

2.5.4 Zprostředkování transakce

IS může zprostředkovat i samotnou transakci sdílení. Tedy umožnit online o sdílení požádat, evidovat dostupnost zařízení (aktuálně i v časovém plánu) a předat informaci o požadavku vlastnickou instituci. V rozhraní pro vlastnickou instituci by pak měly být přehledně

¹⁸ XML, extensible markup language, W3C standard vhodný pro serializaci a výměnu dat, <https://www.w3.org/TR/2006/REC-xml11-20060816/>

¹⁹ RDFa, standard W3C, rozšíření HTML5 o propojená data vkládáním RDF do HTML elementů. <https://www.w3.org/TR/rdfa-primer/>

²⁰ OPD, Organisation Profile Document, <https://opd.data.ac.uk/>, RDF soubor obsahující základní informace o instituci. Definice OPD na <https://opd.data.ac.uk/docs/datasets>.

zobrazeny požadavky na sdílení tak, aby bylo možné organizovat sdílení s ohledem na vnitřní využití vybavení institucí.

Z důvodu zajištění dokumentovatelnosti je potřeba evidovat realizované transakce.

2.5.5 Statistické výstupy

Dalším možným požadavkem na IS může být zpřístupnění dat generovaných používáním systému. Tato data mohou být dále využita jak specifickými nástroji na straně členských institucí nebo organizátora sdílení, ale i nezávisle pro vědecké nebo studijní účely.

Při zpřístupnění dat je potřeba zhodnotit charakter dat a jejich případný neveřejný charakter, určitá skupina dat tak může být přístupná pouze omezenému okruhu uživatelů. Zatímco obsah katalogu lze zpřístupnit bez omezení, například “návštěvnost” jednotlivých zařízení již může být považována za neveřejný údaj, stejně jako transakční historie.

Pro benchmarking (sledování výkonnosti) je nutné mít k dispozici statistické výstupy, ty by měly být k dispozici jak organizátorovi sdílení, tak členským institucím. Statisticky lze zpracovat jak data o obsahu katalogu, tak o uživatelských interakcích s katalogem a také informace o utilizaci jednotlivých přístrojů v rámci sdílení (pokud jsou tato data k dispozici).

3 Případová studie: Equipment.data.ac.uk

V praktické části zkoumaný projekt Equipment.data.ac.uk je celostátní síť v rámci Spojeného království Velké Británie a Severního Irska. Pro účely této případové studie byl zvolen pro svou otevřenost a tedy dostupnost dokumentace i veškerých potřebných dat. Dalším důvodem pro tuto volbu bylo inovativní řešení s využitím propojených dat²¹.

3.1 Metodika

Předkládaná případová studie si klade za cíl zodpovězení otázek:

- Jakým způsobem otevřený katalog Equipment.data.ac.uk podporuje sdílení vědeckého vybavení?
- Jak je řešen sběr dat od členských organizací?
- Jaké procesy, funkce, datové formáty byly využity?
- Jaké je role propojených dat v tomto projektu?
- Jak se vyvíjel počet záznamů (vybavení a členských institucí) v průběhu času?
- Jak se vyvíjel počet členských institucí v průběhu času?
- Nakolik je obsah dat poskytovaných členskými organizacemi validní ve vztahu ke stanovenému standardu (UNIQUIP)?
- Mají údaje validní z hlediska UNIQUIP správnou formu z hlediska jejich předpokládaného obsahu?
- Jak je řešena předmětová kategorizace přístrojů nabízených v katalogu?
- Jaký je poměr zastoupení položek typu facility a equipment?
- Jaké jsou silné a slabé stránky projektu?

²¹ Původní záměr využití propojených dat v projektu formou OPD pro autodiscovery byl nakonec degradován, jak bude zmíněno dále v textu.

3.2 Metody sběru dat

Pro potřeby této případové studie byla získána data z těchto zdrojů:

- veřejně přístupná data na webu projektu:
 - <https://equipment.data.ac.uk/>
 - <https://equipment.data.ac.uk/about>
 - <https://equipment.data.ac.uk/faq>
 - <https://equipment.data.ac.uk/uniqueip>
- opakovaně sledovaný údaj o počtu záznamů a počtu členských institucí na stránce <https://equipment.data.ac.uk/>
- informace o projektu v článcích a dalších relevantních zdrojích:
 - COX, A. a Ch. Gutteridge. equipment.data – Delivering a data autodiscovery infrastructure. CRIS2016: 13th International Conference on Current Research Information Systems (St Andrews, June 9-11, 2016). [cit. 2019-05-24]. Dostupné z <http://hdl.handle.net/11366/50>
 - Data Autodiscovery—The Role of the OPD. *ISPRS International Journal of Geo-Information* [online]. 2016, 5(10), 167 [cit. 2020-01-14]. ISSN 2220-9964.
 - GUTTERIDGE, Christopher J. *Launch of the first HE equipment-sharing database*. Online. In: UNIVERSITY OF SOUTHAMPTON. School of Electronics and Computer Science. 2023. Dostupné z: <https://www.ecs.soton.ac.uk/news/4216>. [cit. 2023-12-17].
 - GUTTERIDGE, Christopher. *Jisc Digifest 2023*. Online. In: SOUTHAMPTON WEB AND DATA INNOVATION TEAM. <https://blog.soton.ac.uk/>. 2023. Dostupné z: <https://blog.soton.ac.uk/webteam/2023/03/08/jisc-digifest-2023/>. [cit. 2023-12-17].
- vstupní data jednotlivých členských organizací stažená ze stránky <https://equipment.data.ac.uk/institutions> dne 12. 12. 2023, která byla dále agregovaná, konvertovaná a analyzovaná php skriptem vytvořeným pro účely této případové

studie²². Vypuštěna data Heriot-Watt University, University of Brighton, Cranfield University, The Open University²³.

3.3 O projektu Equipment.Data

Projekt Equipment.Data má motto “*Enabling access to UK HE research equipment*” - umožňujeme přístup k britskému vysokoškolskému vědeckému vybavení. Za vývojem projektu stojí University of Southampton, financován byl z prostředků EPSRC²⁴, provozovatelem je JISC²⁵[\[18\]](#).

Cílem bylo vytvoření “One stop shop” (vše pod jednou střechou) pro sdílení vědeckého vybavení, které nabízejí členské organizace (vysokoškolské subjekty Spojeného království). Důležitou komponentou je sběr dat od členských organizací a jejich agregace. Výstup je pak obsahem katalogu, ale pro kontrolní a analytické účely je k dispozici ke stažení v různých formátech. Vývoj této služby byl financován v reakci na potřebu zlepšit povědomí o existenci vědeckého vybavení v UK a zvýšit jeho využití. V podmínkách UKRI²⁶ grantů je stanoven požadavek na zadání veškerého vybavení s hodnotou nad 138000 GBP právě do národní databáze ED[\[18\]](#).

Účast v projektu je dobrovolná, pouze organizace čerpající grant UKRI jsou povinné při nákupu přístroje s cenou nad 138000 GBP nahlásit přístroje do katalogu Equipment.data.ac.uk [\[19\]](#).

Principiálně se jedná o katalog, nezajišťuje tedy podporu vlastního sdílení jako žádosti o

22 Důvodem byla od června 2023 nefunkční exportní funkce na stránkách projektu, která nebyla navzdory upozornění a příslibu opravy zprovozněna do data dokončení této práce.

23 Důvodem je nedostupnost vstupních dat, podrobněji v oddílu [Obsah katalogu Equipment.data.ac.uk](#).

24 EPSRC, Engineering and Physical Sciences Research Council, <https://www.ukri.org/councils/epsrc/>

25 JISC, Joint Information Systems Committee, <https://jisc.ac.uk/>, nezisková organizace zajišťující ve Spojeném království internetovou infrastrukturu a některé nad tím vybudované služby vysokoškolským organizacím.

26 UKRI, UK Research and Innovation, <https://www.ukri.org>, organizace financovaná z Department for Science, Innovation and Technology (DSIT), která zastřešuje mj. všech 7 Research Councils, státních agentur pro kompetitivní financování výzkumu formou grantů.

rezervace, přehled aktuální dostupnosti a podobně.

Přínosem projektu je tedy možnost zjistit, zda je potřebné vybavení k dispozici v síti členských subjektů. Pochopitelně může výzkumníkovi pomoci i při vyhledávání v rámci své vlastní organizace.

Vlastní agenda kolem sdílení pak již probíhá konvenční cestou (e-mail, telefonicky).

3.4 Historie projektu

V roce 2008 vyvinula Materials Research School and the Centre for Engineering and Design Education při Univerzitě v Loughborough „Equipment Database“ (databázi přístrojů), online katalog laboratorního vybavení, dílenských strojů a speciálního vybavení z celé univerzity. Tento katalog, který nyní obsahuje několik tisíc položek, umožňuje zaměstnancům a studentům vyhledat konkrétní položku a tyto si půjčovat, rezervovat nebo najímat pro výzkumné nebo výukové účely.

V březnu 2011 JISC financovala další vývoj včetně zpřístupnění veřejnosti a využití otevřených propojených dat. Vylepšená aplikace Kit-catalogue™ je k dispozici jako software s otevřeným zdrojovým kódem (<https://www.kit-catalogue.com/>) od prosince 2011.

Získané zkušenosti byly dále využity v projektu Uniquip²⁷, který definoval standardy pro zveřejňování údajů o vědeckém vybavení[20].

Následoval vývoj equipment.data.ac.uk. Ten byl financován EPSRC. Cílem byl jednoduchý a snadno reprodukovatelný nástroj umožňující získávání a agregaci dat z databází vědeckého vybavení britských výzkumných institucí do jediného prohledávatelného portálu.

Vývoj equipment.data byl založen na partnerství mezi řadou univerzit ve Spojeném

²⁷ Nyní přístupné na <https://www.uniquip.ecs.soton.ac.uk/>, původní doména <http://www.uniquip.org> již není dostupná.

království. Základem byl výše zmíněný projekt UNIQUIP, který definoval standardy pro zveřejňování údajů o vědeckém vybavení[18][21].

Projekt byl spuštěn v dubnu 2013[13].

Změny mezi roky 2022 a 2023:

- Zrušen export v RDF, nyní již jen CSV²⁸ a JSON²⁹.
- Zrušen ingest status (přehled vstupních dat, prakticky nahrazeno funkcí Institutions).
- Přidána stránka Institutions, kde je přehled dat jednotlivých institucí.
- Importní formát omezen na CSV a XLS/XLSX³⁰. RDF opuštěno.
- Zrušen harvesting dle OPD (info se posílá prostřednictvím online formuláře, změny je nutno hlásit stejnou cestou).

3.5 Procesy

3.5.1 Katalog

Katalog zobrazuje informace o vědeckém vybavení, které je k dispozici ke sdílení.

Vyhledávání probíhá v údajích:

- Name (název vybavení)
- Manufacturer (výrobce)
- Model (model výrobku)
- Description (popis vybavení volným textem)

Ověřeno pokusem o vyhledávání v ostatních údajích dle UNIQUIP podle hodnot v náhodně

28 CSV neboli comma separated values - je textový soubor, kde jednotlivé hodnoty odděluje separátor, v tomto případě znak čárky (ASCII 44). Jedná se o běžný formát pro publikování jednoduše strukturovaných dat. Tabulkové procesory je dokáží otevřít. Je založený na standardu RFC4180.

29 JSON, JavaScript Object Notation, <https://ecma-international.org/publications-and-standards/standards/ecma-404/>. Multiplatformní způsob zápisu dat (datový formát), určený pro přenos dat, která mohou být organizována v polích nebo agregována v objektech. Standard byl zaveden Ecma International – European association for standardizing information and communication systems, <https://ecma-international.org/>.

30 Formáty používané aplikací Microsoft Excel. XLS je proprietární binární formát souboru (popis formátu byl později společností Microsoft zveřejněn). XLSX je otevřený formát navržený společností Microsoft. Jedná se o ZIP komprimovaný XML soubor.

vybraných záznamech v agregovaných datech.

Filtrování je možné podle parametrů:

- Institution (instituce, buď jmenovitě, nebo dle lokálního konsorcia)

Údaje u vyhledaného záznamu (detail záznamu):

Zobrazují se údaje dle definice UNIQUIP - vždy jen ty, které jsou v záznamu přítomné:

- Name (název vybavení)
- Manufacturer (výrobce - sloučeno do Description)
- Model (model - sloučeno do Description)
- Description (popis vybavení volným textem, popisu je předřazen výrobce a model, pokud je přítomen v údajích Manufacturer a Model)
- Related Facility ID (ID nadřazeného zařízení, pokud je přístroj součástí zařízení)
- Technique (kategorizace oboru, technologie)
- Location
- Photo
- Department
- Site Location
- Building
- Service Level
- Web Address
- Následují kontaktní informace (primární a sekundární) dle UNIQUIP standardu.

Ověřeno zobrazením náhodně vybraných záznamů.

3.5.2 Předávání dat

Registrace nového člena konsorcia: Členské organizace musí nejprve poskytnout základní informace o sobě. Jde např. o název organizace, logo, adresu webové stránky, přibližné souřadnice, kontakt pro opravy, ale zejména umístění, ze kterého lze sklízet data vybavení (asset data). Tyto informace byly součástí OPD (viz pododdíl Autodiscovery), v současné době se zadávají (včetně změn těchto údajů) prostřednictvím online formuláře Institutions

Submission (<https://forms.office.com/e/SKAEHrjvW4>).

Vystavení dat členem konsorcia: Člen konsorcia vystavuje data v požadované podobě (podrobněji v oddíle [Data na vstupu](#)) na svých webových stránkách, nebo ve službě cloudového úložiště. Zdrojem dat pro vystavovaný soubor může být ruční zadání (k dispozici je vzorový soubor tabulkového procesoru, tento lze ručně vyplnit údaji o vybavení) Také je možné exportovat data o vybavení ze stávajících informačních systémů (např. CRIS).

Harvesting: Sběr dat členských organizací probíhá periodickým stahováním souborů vystavených organizacemi (tzv. harvesting = sklizeň).

Dokumentace uvádí denní interval sklizně[22], pozorováním stránky Institutions (<https://equipment.data.ac.uk/institutions>), sloupec “Date harvested” se ale ukazuje, že interval je v praxi delší, přibližně týdenní.

3.5.3 Autodiscovery

Aby mohla být sklizena data členských organizací, musí agregátor vědět, kde má data hledat. K tomuto účelu slouží v případě projektu Equipment.data.ac.uk tzv. OPD - Organisation Profile Document, RDF soubor obsahující základní informace o instituci včetně odkazů na datové zdroje, tedy i zdroj dat o sdíleném vědeckém vybavení pro harvesting[14].

Odkaz na OPD je umístěn v html hlavičce domovské stránky instituce[14]:

```
<link rel="openorg" href="http://www.example.ac.uk/profile.ttl" />
```

V případě, že instituce z nějakého důvodu nemá možnost měnit hlavičku domovské stránky, může použít metodu “well-known”. V takovém případě umístí OPD do umístění “/.well-known/openorg”, tedy např. v případě domény cuni.cz bude cesta k OPD “<https://www.cuni.cz/.well-known/openorg>”[14].

Profil je pravidelně ověřován agregátorem. Výhodou tohoto řešení je, že jakékoliv změny

informací o instituci (včetně umístění souborů pro sklizení dat do katalogu) se promítnou při nejbližší sklizni automaticky, bez nutnosti lidského zásahu. O celém procesu zjišťování umístění dat na základě OPD pak autoři hovoří jako o Autodiscovery procesu[14]. Jedná se o ukázkové využití linked open data.

Přehled všech sklizených OPD dat je k dispozici na webu <https://opd.data.ac.uk/>.

Definice OPD je dostupná na <https://opd.data.ac.uk/docs/datasets>.

Aktualizace k červenci 2023: Nové organizace již nejsou instruovány k tvorbě OPD a vkládání linku na OPD do domovských stránek. Namísto toho je požadováno, aby nové organizace (a stávající v případě aktualizace informací) vyplnili formulář Institutions Submission form[22]. Lze tedy předpokládat, že v tuto chvíli se již autodiscovery na základě OPD nepoužívá. Tuto hypotézu potvrzuje i fakt, že v OPD registru se v současnosti (údaj aktualizován k 12. 12. 2023) nachází jen 21 institucí[23]. zatímco konsorcium Equipment.data.ac.uk má ve stejném okamžiku 59 členů[24].

3.5.4 Kontrolní výstupy

Kontrolní výstupy na stránce Institutions: Přehled je generován formou tabulky, každou instituci reprezentuje jeden řádek. Je uveden:

- název instituce
- ROR identifikátor³¹ instituce
- počet záznamů v datasetu instituce
- katalog (vyhledávání v katalogu pouze v rámci dat instituce)
- funkce stažení vstupních dat (Datasource)
- okamžik posledního pokusu o sklizeň

³¹ ROR, Research Organization Registry, <https://ror.org/> je otevřený registr výzkumných organizací. V současnosti zahrnuje jedinečné identifikátory a metadata více než 102000 výzkumných organizací.

- funkce kontroly chyb (Data Check) zobrazuje statistické údaje o sklizních:
 - Graf aktivity (počty sklizených záznamů v časové ose)
 - Okamžik poslední sklizně
 - Stav sloupců (chybějící nepovinné a povinné údaje, údaje navíc (ve vztahu k definici UNIQUIP))
 - Výpis čísel řádků s chybami (ve smyslu předchozího bodu)

3.6 Datové formáty

Data projektu lze rozdělit na dvě zásadní skupiny:

- **Data na vstupu**, tedy ta, která členské organizace vystavují a jsou stahována v rámci procesu harvestingu (sklizeně).

Data na vstupu jsou dvojího druhu:

- Data o organizaci (OPD, organisation profile data)
 - Data vybavení (asset data)
- **Data na výstupu**, tedy data, která jsou výsledkem procesu autodiscovery, agregována a vystavena na webu equipment.data.ac.uk (a použita pro vlastní funkci katalogu). Ta jsou určena ke kontrolním a analytickým účelům, díky otevřenosti systému však mohou sloužit i jako zdroj pro alternativní vyhledávač.

3.6.1 Data na vstupu

Data o organizaci byla původně sklizena z OPD (zmíněno v oddíle [Procesy](#)). V současné době se data o organizaci získávají vyplněním online formuláře Institutions Submission (<https://forms.office.com/e/SKAEHrjvW4>).

Vlastní data o vybavení (asset data) mohou organizace vystavit na své webové stránce nebo pomocí libovolné cloudové služby (např. Google Docs), pokud není přístup k souboru omezen[22] (soubor musí být volně dostupný, což některé instituce nedodržely, podrobněji uvedeno v oddílu [Obsah katalogu](#)).

Takto umístěná data jsou pravidelně sklízena do databáze katalogu[22]. Proces sklizně i řízení zdroje sklizně procesem autodiscovery je popsán v oddíle [Procesy](#).

Data se ukládají ve strojově čitelné formě jako soubory CSV³² (preferovaná varianta) nebo XLS/XLSX³³. Jednotlivé údaje v tabulce musí odpovídat standardu UNIQUIP. Nezbytné je používat přesné názvy záhlaví sloupců v Uniquip kvůli správnému přiřazování k údajům dle UNIQUIP[22].

Datový formát UNIQUIP byl navržen tak, aby bylo snadné jej generovat a obsahoval minimální množství požadovaných informací. Byl vytvořen spoluprací zástupců z 30 britských univerzit a jeho cílem bylo vytvořit nejzákladnější formát, který by usnadnil sdílení vybavení ve Spojeném království. Stručně řečeno, každá položka musí mít popis, umístění a nějaký druh kontaktních informací. Vše ostatní je volitelné[22].

Definice formátu UNIQUIP je vložena jako příloha 2.

Přestože dokumentace nabízí pouze jediný datový formát UNIQUIP v podobě tří možných typů souborů (csv, xls, xlsx), přehled na stránce Institutions (<https://equipment.data.ac.uk/institutions>) ukazuje i další formáty, které jsou používány. Systém je tedy zjevně umí načíst a sloučit tak, aby odpovídaly UNIQUIP formátu:

³² CSV neboli comma separated values - je textový soubor, kde jednotlivé hodnoty odděluje separátor, v tomto případě znak čárky (ASCII 44). Jedná se o běžný formát pro publikování jednoduše strukturovaných dat. Tabulkové procesory je dokáží otevřít. Je založený na standardu RFC4180.

³³ Formáty používané aplikací Microsoft Excel. XLS je proprietární binární formát souboru (formát byl později společností Microsoft otevřen). XLSX je otevřený formát navržený společností Microsoft. Jedná se o ZIP komprimovaný XML soubor.

- Formát KitCat³⁴ (nyní využívá šest organizací: Loughborough University, The Open University, The University of Bristol, The University of Nottingham, University College London (UCL), University of Leicester). Sklizen v JSON³⁵ souborech.
- Formát PURE5³⁶ (nyní využívá University Of Brighton)

Z textu bodu 13 vstupního formuláře[25] vyplývá, že organizace nově se zapojující do projektu mohou poskytovat již jen CSV soubory ve formátu UNIQUIP, zatímco stávající mohou data poskytovat i z jiných platforem (Kit-Catalogue). Tyto instituce ale budou kontaktovány s žádostí o změnu (myšleno na CSV ve formátu UNIQUIP) za účelem zlepšení kvality poskytovaných dat.

To je také důvodem, proč formáty KitCat a PURE5 nejsou zmiňovány v dokumentaci. Přestože je zachována určitá zpětná kompatibilita pro organizace využívající tyto formáty, je zde evidentní záměr sjednocení na UNIQUIP.

3.6.2 Data na výstupu

Členské instituce i veřejnost mají otevřený přístup k agregovaným datům za celý projekt. Tyto mohou sloužit ke kontrole procesu importu z vystavených souborů, ale také mohou být použita pro sofistikované vyhledávání nad rámec schopností katalogu Equipment.data.ac.uk.

Kombinovaná (agregovaná) data jsou k dispozici v těchto formátech[24]:

- CSV (lze otevřít v tabulkovém procesoru)

³⁴ Formát používaný katalogem Kit-Catalogue, provozovaným Loughborough University (<https://www.kit-catalogue.com/projectpages/>). Jde o open source katalog vědeckého vybavení, postavený na PHP/MySQL, předchůdce Equipment.data.ac.uk. Informace o údajích v katalogu je uvedena v dokumentaci (https://www.kit-catalogue.com/projectpages/wiki/index.php/An_introduction_to_Kit-Catalogue#An_overview_of_the_key_features).

³⁵ JSON, JavaScript Object Notation, <https://ecma-international.org/publications-and-standards/standards/ecma-404/>. Multiplatformní způsob zápisu dat (datový formát), určený pro přenos dat, která mohou být organizována v polích nebo agregována v objektech. Standard byl zaveden Ecma International – European association for standardizing information and communication systems, <https://ecma-international.org/>.

³⁶ Pure je CRIS/RIMS systém společnosti Elsevier.

- JSON

Kombinovaný výstup v CSV a JSON koresponduje se vstupním standardem Uniquip. Je zde však odlišnost u institucí využívajících jako vstupní formát dat KitCat: ve výstupních kombinovaných datech se vyskytují údaje nad rámec definice Uniquip, a to “Manufacturer” a “Model”. Jde tedy o údaje (velmi přínosné), přítomné v KitCat a nedefinované v Uniquip, které tvůrci katalogu rozhodli přenést do kombinovaného výstupu.

Tato kombinovaná data obsahují navíc ještě informace převzaté z OPD (identifikaci instituce, která přístroj nabízí):

- Institution Name
- Institution URL
- Institution Logo URL
- Datestamp
- Approximate Coordinates
- Corrections

Počínaje vzorkem dat z 11. 5. 2023 v exportovaných datech chybí výše uvedené údaje přebírané z OPD.

Nejpozději od 13. 7. 2023 je tento export nefunkční (poslední známé datum, kdy byl tento export funkční, je 10. 5. 2023 a 15. 6. 2023 byla chyba zjištěna).

- Dne 15. 6. 2023 byla kontaktována podpora (Jisc Service Desk) s upozorněním na chybu a žádostí o opravu.
- Dne 13. 7. 2023 byla ze strany Jisc Service Desk obdržena odpověď (Ref:IN:01395820): *Yes we have an issue at the moment with our system. Once this is fixed I will email you back.*
- Ke dni odevzdání této práce **nebyl export opraven**, ani nepřišla jiná reakce

Laxní přístup podpory k výpadku jedné z funkcí katalogu bohužel ukazuje na fakt, že se projekt patrně dostal do terminálního stadia.

Do roku 2021 byl ke stažení také tzv. ingest status (ve formátu JSON), který obsahoval vstupní data od členských institucí. Náhradou ingest statusu je nyní stránka Institutions (<https://equipment.data.ac.uk/institutions>), kde je ve sloupci “Datasource” ke stažení vstupní sada dat každé instituce.

3.7 Otevřená propojená data

Praktickou aplikací otevřených propojených dat v projektu Equipment.data.ac.uk mělo být využití OPD pro proces Autodiscovery. Jak již bylo zmíněno v pododdíle Autodiscovery, tento záměr byl nakonec opuštěn. Možnými důvody se práce zabývá v diskuzi.

3.8 Obsah katalogu Equipment.data.ac.uk

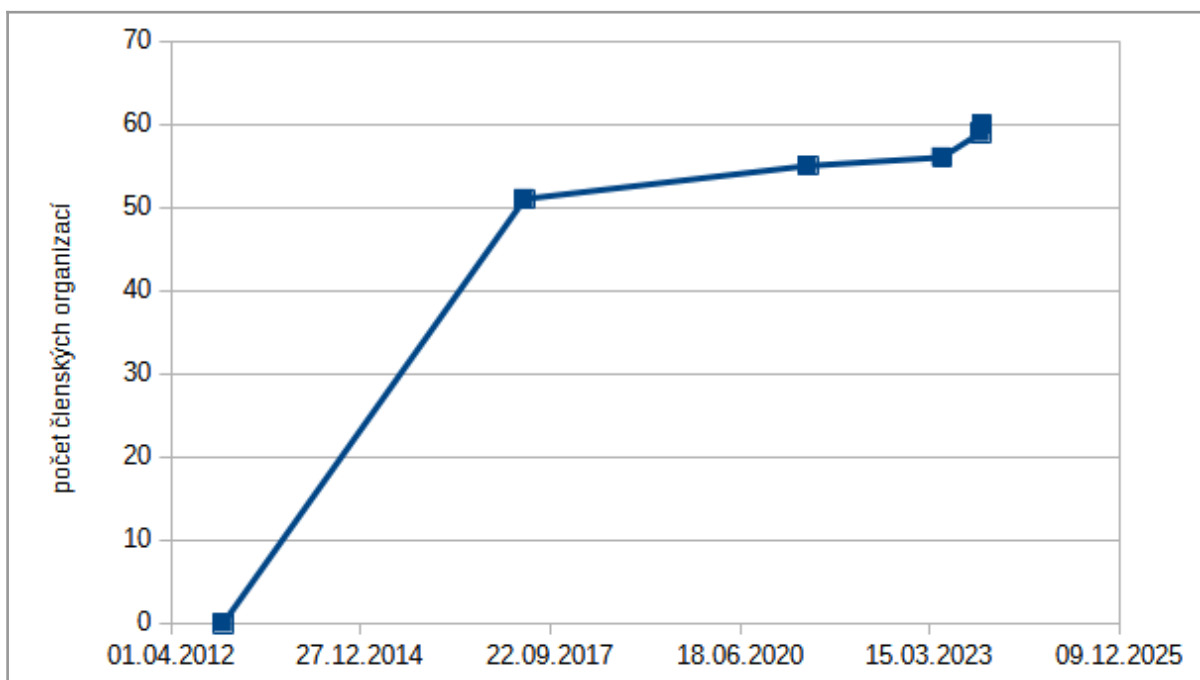
3.8.1 Databáze, počet záznamů a průběh v čase

Vývoj počtu zapojených organizací a položek v katalogu je prezentován v Tabulce 1.

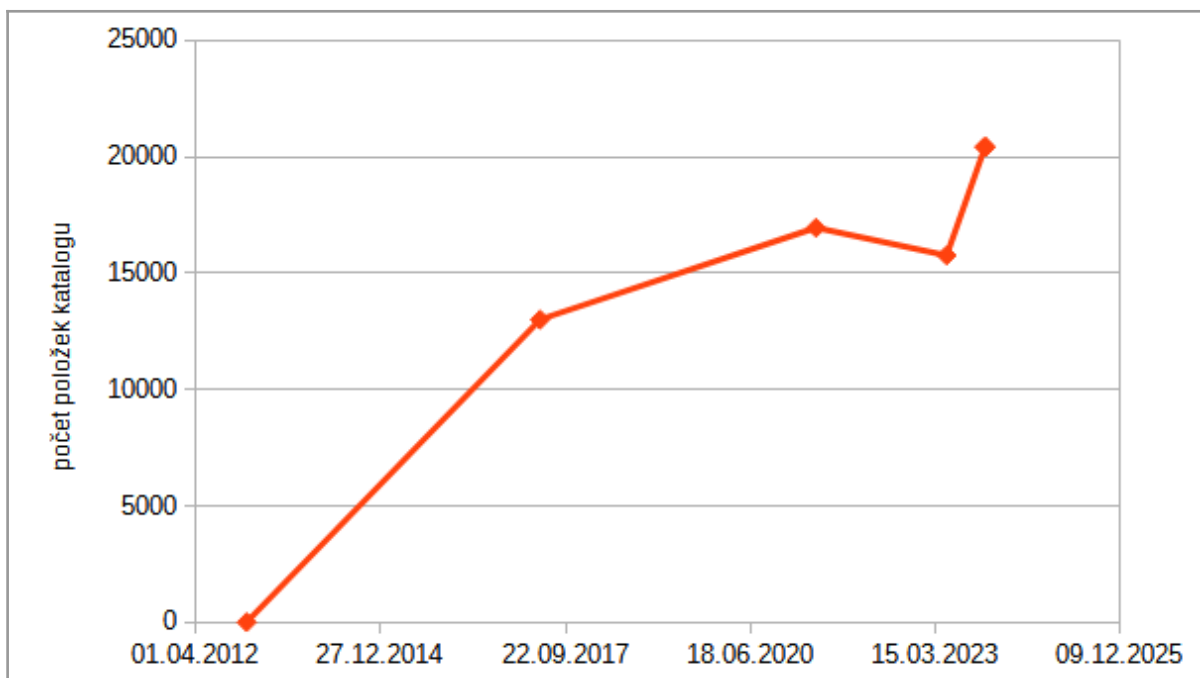
datum	počet organizací	počet položek
8. 5. 2017	51	>13 tisíc
10. 6. 2021	55	16 945
18. 5. 2023	56	15 757
8. 12. 2023	59	20 417
16. 12. 2023	60	20 420

Tabulka 1: Vývoj počtu zapojených organizací a počtu položek v katalogu equipment.data.ac.uk.

Zdroj: Pravidelné sledování stavové informace na[24], údaje za rok 2017 získány z[26].



Graf 1: Vývoj počtu členských organizací.



Graf 2: Vývoj počtu záznamů vybavení.

3.8.2 Informace o vzorku dat

V následující analýze dat je použit vzorek dat z 12. 12. 2023. V této době již nebyl export agregovaných dat funkční, jak je uvedeno v oddíle [Data na výstupu](#).

Z důvodu nefunkčnosti exportu agregovaných dat (a se záměrem analyzovat aktuální data) byl vytvořen skript v PHP pro čtení a zpracování dílčích vstupních dat jednotlivých organizací. Skript je pro účely dokumentace uložen v otevřeném repozitáři Zenodo: <https://doi.org/10.5281/zenodo.10396717>, stejně jako finální vzorek dat.

Finální vzorek dat byl získán ze vstupních dat jednotlivých organizací na stránce Institutions³⁷ výše uvedeným skriptem a následnou agregací do jedné sady dat. V případě dat uložených ve formátu KitCat (json) byla provedena konverze do polí dle standardu UNQUIP přes mapovací tabulku, která byla také uložena v otevřeném repozitáři Zenodo: <https://doi.org/10.5281/zenodo.10396717>. V případě souborů uložených v XLSX formátu byla provedeno pomocí LibreOffice Calc přeuložení do CSV formátu s nastavením čárky jako oddělovače a s uzavřením hodnot do uvozovek dle potřeby.

Dalším přínosem tohoto postupu je, že pracujeme se syrovými vstupními daty a výzkum tedy není ovlivněn případným processingem během agregace katalogem.

Tabulka Institutions obsahovala v době načtení vzorku 59 institucí.

Do tohoto vzorku dat nebyla zahrnuta data organizací, které navzdory stanoveným podmínkám³⁸ svá data nezpřístupnily otevřeně, ale přístup chrání potřebou přihlásit se (Heriot-Watt University), API klíčem³⁹ (University of Brighton) nebo jejich vstupní data

³⁷ <https://equipment.data.ac.uk/institutions>

³⁸ The URL MUST be persistent and open (**not behind any login**). <https://equipment.data.ac.uk/faq#faq11>, step 3, třetí odrážka.

³⁹ Organizace byla kontaktována se žádostí o sdělení API klíče a prosbou o vysvětlení důvodu ochrany. Na dotaz nebyla do data odevzdání práce obdržena odpověď.

nejdou dostupná z důvodu chyby stahování (The Open University), případně jsou data poškozená (Cranfield University, University of Cambridge, University of Glasgow).

Takto bylo vyloučeno celkem šest institucí, do analýzy tedy vstupuje 53 institucí.

Data takto vyloučených institucí nejsou v procesu harvestingu načtena ani do katalogu Equipment.data.ac.uk, v případě poškozených dat ale zůstávají v katalogu záznamy platné před výskytem chyby a počet těchto položek se zobrazuje i v přehledu Institutions (Cranfield University 96 záznamů, University of Cambridge 3866 záznamů, University of Glasgow 63 záznamů). Z hlediska celkového počtu záznamů v katalogu vs. celkového počtu záznamů načtených skriptem je tedy nutné vzít v potaz, že jsou v katalogu starší verze dat těchto tří institucí, které nemají odpovídající záznamy v analyzovaných datech načtených ze vstupních souborů na stránce Institutions. Celkem se jedná o 4025 záznamů.

Dále bylo zjištěno 14 nadbytečných řádků u dat British Ocean Sediment Core Research Facility.

Kontrola: Skriptem načteno 16378 řádků dat, na <https://equipment.data.ac.uk/> uvedeno k 12. 12. 2023 celkem 20417 položek vybavení. $16378 + 14 + 4025 = 20417$.

Pokud jde o souborové formáty, Uniquip je převážně uložen v CSV souboru, pouze čtyři instituce vystavují soubory ve formátu XLSX. Všech pět institucí využívajících datový formát KitCat vystavuje své soubory v souborech JSON.

Z celkového počtu 59 institucí jich 27 použilo k vystavení souboru pro proces harvestingu online dokument služby Google. Ve všech případech se jedná o CSV soubor.

3.8.3 Výsledky analýzy dat

Na výzkumnou otázku “**Nakolik je obsah dat poskytovaných členskými organizacemi**

validní ve vztahu ke stanovenému standardu (UNIQUIP)?” odpovídá Tabulka 2.

povinný údaj	celkem záznamů	přítomno	podíl validních záznamů
Name Description	16378	16374	99,98 %
Primary contact telephone URL e-mail	16378	15610	95,31 %
Secondary contact telephone URL e-mail	16378	8479	51,77 %

Tabulka 2: Validita ve vztahu ke standardu UNIQUIP (přítomnost povinných údajů).

Povinnost vyplnit název nebo popis je vyplněna prakticky bezvýhradně, uvedené čtyři záznamy jsou prázdné řádky v exportních datech. Povinnost vyplnění primárních kontaktů je dodržována. Horší je to s vyplněním sekundárních kontaktů, které jsou dle UNIQUIP také povinné.

Na výzkumnou otázku **“Mají údaje validní z hlediska UNIQUIP správnou formu z hlediska jejich předpokládaného obsahu?”** odpovídá Tabulka 3.

údaj	předpokládaný obsah	přítomno	validních záznamů	podíl validních záznamů
location	odkaz na wikipedii	7468	1443	19,32 %
primary contact e-mail	email	12163	10768	88,53 %
primary contact url	url	5840	4972	85,14 %
secondary contact e-mail	email	6172	5983	96,94 %
secondary contact url	url	2371	1664	70,18 %
photo	url ⁴⁰	2608	2435	93,37 %
web address	url	7502	7014	93,5 %

Tabulka 3: Validita ve vztahu k předpokládanému obsahu údaje.

40 Nelze aplikovat kontrolu koncovkou souboru, protože obraz může vracet i funkce webové stránky volaná pomocí url. Bylo by nutné kontrolovat obsah vrácený voláním url.

Nejvíce problematickým z hlediska validity hodnoty ve vztahu k předpokládanému obsahu je údaj “location”. Dle definice UNIQUIP má obsahovat odkaz na heslo příslušného místa na Wikipedii (“The URL of a Wikipedia page for the city/town/borough in which the facility is in or near...”)[27].

Tuto podmínku plní jen 19 % zadaných lokací. Většinou údaj obsahuje slovní název místa, případně odkaz do informačního systému instituce na stránku s informacemi o místě.

U ostatních údajů chyby představují spíše chybnou syntaxi (v url často absenci protokolu před “www”), prefix “mailto:” před e-mailem, vyplnění hodnot “0” nebo “N/A” do prázdného údaje nebo pokus vložit do jednoho údaje více URL / e-mailů s použitím oddělovače.

Na výzkumnou otázku “**Jak je řešena předmětová kategorizace přístrojů nabízených v katalogu?**” odpovídá Tabulka 4.

Údaj “Technique” - kategorizace	přítomno	použito	podíl
N8	3595	1961	54.55 %
CPV	3595	0	0 %

Tabulka 4: Použité taxonomie v údaji Technique.

V datech byla prověřena přítomnost termínů dle taxonomie N8 nebo CPV v údaji Technique. Dokumentace neuvádí, zda to má být prefix nebo sufix (“by adding”)[27]. Byly prověřeny vždy obě varianty.

Údaj Technique často nebývá vyplněn (jen 22 % záznamů). V případě vyplnění je přibližně v polovině případů vyplněn volně tvořenými termíny. Pouze University of Leeds používá (exportuje z UNIQUIP) taxonomii N8 (1961 záznamů). Řetězec “:N8” používá jako sufix, uvádí pak jen poslední úroveň hierarchie (druh).

Údaje z číselníku CPV se v poli Technique ve sledovaných datech nenacházejí.

Celkově lze tedy využití kategorizace vybavení v katalogu označit jako sporadické.

Na výzkumnou otázku “**Jaký je poměr zastoupení položek typu facility a equipment?**” odpovídá Tabulka 5.

Údaj “type”	přítomno	použito	podíl
Equipment	12257	11758	95,93 %
Facility	12257	319	2,6 %
jiné hodnoty	12257	180	1,47 %

Tabulka 5: Facility vs. Equipment.

Naprostá většina vybavení v katalogu je typu “equipment”, tedy spíše samostatné přístroje. Údaj, přestože nepovinný, je do převážně vyplněn ($\frac{3}{4}$ záznamů). Hodnoty označené “jiné hodnoty” jsou zpravidla chyby při psaní, volně psané termíny označující typ vybavení (mimo smysl rozdělení equipment / facility) a v některých případech použití výrazu “facility” v množném čísle”.

3.9 Silné a slabé stránky, perspektivy

Silné stránky, přínosy:

- Katalog nepochybně usnadnil sdílení vědeckého vybavení v rámci britské akademické sféry. Základní funkci, tedy možnost vyhledat přístroj, zjistit, zda je k dispozici a získat kontaktní údaje, katalog plní.
- Zadání dat do katalogu probíhá s minimálními nároky na členské instituce (žádné speciální rozhraní / aplikace, pouze vystavení souboru tabulkového procesoru).

Slabé stránky, nedostatky:

- Chybí podpora (administrace) vlastního procesu sdílení (tedy dostupnost zařízení, volné termíny, možnost rezervace. Tato podpora by usnadnila celý proces a archivace dat z těchto procesů by umožnila analyzovat využití jednotlivých přístrojů.
- Chybí kategorizace (sporadické využití taxonomie N8 u jedné organizace). Jednotlivé přístroje nejsou nijak kategorizovány, až na výjimky (KitCat) není rozdělen výrobce a model do samostatných údajů. To komplikuje vyhledávání a činí celý katalog

nepřehledným. Kategorizace není neřešitelná, jak ukazuje projekt konsorcia N8.

- I v případě kategorizace v datech (University of Leeds) není tato v katalogu nijak využita, katalog ji pouze zobrazí jako údaj v detailu přístroje. Nelze podle ní vyhledávat nebo procházet, nelze zobrazit hierarchický strom.
- Kvalita dat katalogu (chybějící povinné údaje, vyplnění údajů v jiné než očekávané podobě, nefunkční odkazy). Podrobněji v pododdílu [Výsledky analýzy dat](#).
- Poměrně primitivní uživatelské rozhraní katalogu (nemožnost volit prohledávaný údaj, omezené možnosti filtrování, zobrazení slučuje údaje Manufacturer a Model do popisu).
- Není žádná nápověda ke katalogu, není zřejmé, ve kterých údajích bude hledáno.

Perspektivy:

Lze konstatovat postupný úpadek systému:

- redukce možných formátů vstupních i výstupních dat
- opuštění konceptu propojených dat (autodiscovery pomocí OPD) a přechod na online formulář
- snížení četnosti sklizně dat
- nefunkční agregovaný export (a ignorování tohoto problému po dobu měsíců)
- slabá uživatelská podpora

Tento úpadek konstatuje i sám autor projektu v příspěvku z konference Jisc Digifest 2023. Jak autor uvádí, po předání projektu z univerzity v Southamptonu do rukou Jisc nebyl dodržen slib o zachování údržby a vývoje a systém z pohledu autora degraduje[28].

S ohledem na zjevně upadající podporu projektu lze očekávat spíše stagnaci, kdy bude projekt zachován do doby, než jej nahradí jiný, bez dalšího vývoje. Jako pravděpodobný scénář lze označit pokračování úpadku s postupným omezením na nutné minimum z hlediska nároků na podporu a údržbu. Nelze však ani vyloučit jiné scénáře, včetně úplného uzavření bez náhrady, nebo naopak oživení zájmu a přísunu zdrojů.

4 Srovnání s obdobnými projekty

4.1 Obdobné projekty v zahraničí

Ve **Velké Británii** existuje několik dalších **regionálních konsorcií** pro sdílení vědeckého vybavení[22]:

- GW4 <https://gw4.ac.uk/>
- Midlands Innovation <https://midlandsinnovation.org.uk/>
- N8 Research Partnership <https://www.n8research.org.uk/>
- Science and Engineering South (SES) <https://www.ses.ac.uk/>
- Wales Innovation Network <https://uniswales.ac.uk/our-work/wales-innovation-network>
- Yorkshire Universities <https://yorkshireuniversities.ac.uk/>

Z uvedených konsorcií zasluží podrobnější zmínku konsorcium **N8 Research Partnership** (konsorcium osmi významných univerzit ze severu Velké Británie: Durham University, Lancaster University, University of Leeds, University of Liverpool, University of Manchester, Newcastle University, University of Sheffield, University of York), které ve svém katalogu úspěšně používá vlastní kategorizaci výzkumného vybavení, postavenou na tříúrovňové hierarchii (Class → Order → Genus, tedy Třída → Řád → Druh).

V Evropě byla spuštěna celoevropská síť pro sdílení vědeckého vybavení v rámci EOSC marketplace <https://marketplace.eosc-portal.eu/>. V současné době ještě neobsahuje data, pravděpodobně ještě potenciální uživatelé nejsou informováni o této možnosti.

Z aktivních celoevropských projektů lze zmínit například **Euro-BioImaging** <https://www.eurobioimaging.eu/about-us/about-eubi>, což je výzkumná infrastruktura zaměřená na zobrazovací metody v biologii a biomedicíně, vzniklá za podpory European Strategy Forum on Research Infrastructures (ESFRI) <https://esfri.eu/>. Tři z uzlů Euro-BioImaging jsou v České Republice (Akademie věd ČR, Univerzita Karlova a Masarykova Univerzita v Brně)[29].

I tento projekt využívá kategorizaci vybavení v tříúrovňové hierarchii, v tomto případě hraje kategorizace hlavní roli, neboť výhradním způsobem k vyhledání vybavení je právě procházení jednotlivých úrovní hierarchie. Po dosažení nejnižší úrovně je ještě nabídnuta volba konkrétního uzlu (země, lokality) a následně je možné podat návrh (na realizaci projektu na daném zařízení).

Euro-BioImaging mimo poskytnutí zařízení nabízí i školicí služby a služby zpracování dat produkovaných zobrazovacími metodami.

V USA lze uvést jako příklad **Materials Research Facilities Network (MRFN)** <https://mrfn.org/>, což je celostátní síť sdílených výzkumných zařízení (Shared Experimental Facilities) v oblasti výzkumu materiálů podporovaná Material Research Science and Engineering Centers <http://www.mrsec.org/> v rámci National Science Foundation <http://www.nsf.gov/>. Jejich katalog propojuje 19 univerzit, 30 center a 143 zařízení (facility). Obsahuje také nabídku spolupráce s 230 odborníky. Nabízen je i přístup k vzdělávacím a výukovým materiálům⁴¹[30].

Obdobně jako N8 Research Partnership i katalog MRFN obsahuje kategorizaci v tříúrovňové hierarchii. Touto kategorizací lze filtrovat výsledky vyhledávání nebo přímo procházet hierarchickým stromem.

V Asii je sdílení vědeckého vybavení věnována velká pozornost, čemuž odpovídá i výsledek rešerše, kdy nejvíce článků k dané problematice pochází právě z této oblasti. Konkrétním příkladem může být jihokorejský **ZEUS** (Zone for Equipment Utilization Service) <http://www.zeus.go.kr>, portál pro sdílení vědeckého vybavení. V tomto případě se nejedná pouze o katalog, ale i o jednotný rezervační systém pro všechny členské organizace. V systému se zároveň provádí i zadávání nového vybavení a integrovaný management výzkumných zařízení a nákladů.

41 Ten je však poměrně chudý (tři dokumenty).

Jde o jednu ze služeb NFEC (National Research Facility Promotion Center) <https://www.nfec.go.kr/> - organizace pro podporu výzkumných zařízení. Tato specializovaná organizace mimo služby ZEUS poskytuje i služby RED (pro posuzování a plánování nákupů výzkumného vybavení) a FEEL (pro přístup k reportům výzkumných zařízení). Všechny tyto služby jsou propojeny, takže i nákup vybavení je centrálně posuzován, organizován a plánován, což pomáhá efektivnímu využití prostředků na nákup vědeckého vybavení[10].

Přehled uvedených projektů je uveden v Tabulce 6.

projekt	založeno	stav	region	obor	počet členů	počet položek ⁴²	kategorizace
EQ.D ⁴³	2013	aktivní	UK	všeoborové	59	20417	NE ⁴⁴
N8	2016 ⁴⁵	aktivní	UK	všeoborové	8	3849	ANO, 3U ⁴⁶
MRFN	-	aktivní	USA	materiálové vědy	19	1240	ANO, 3U
EOSC Marketplace	2020 ⁴⁷	spuštěno, nenaplněno	EU	všeoborové	-	-	-
EUBI	2020	aktivní	EU	zobrazovací metody v biologii a biomedicině	35	110	ANO, 3U
ZEUS	2013	aktivní	Jižní Korea	všeoborové	2810	54486 ⁴⁸	-

Tabulka 6: Srovnání vybraných projektů sdílení vědeckého vybavení diskutovaných v oddílu 4.1

4.2 Obdobné projekty v České republice

České vysoké učení technické (ČVUT), <https://www.cvut.cz/> provozovalo databázi DNEP (Databáze nabídek, expertů a přístrojů ČVUT).

42 Údaje aktualizovány k 14. 12. 2023

43 Equipment.data.ac.uk

44 Načteno v datech jedné z organizací, bez využití v katalogu.

45 Údaj dle copyrightu, nelze ověřit.

46 3U = tři úrovně hierarchie

47 <https://eoscssecretariat.eu/news-opinion/eosc-portal-unveils-new-release>

48 Počet zařízení v režimu sdíleného využití, pouze sektor výzkum a vývoj (R&D).

Přístrojové vybavení však bylo reprezentováno jen malým vzorkem (desítky přístrojů).

Záznam přístroje obsahoval následující údaje:

- název česky
- název anglicky
- kategorie
- datum zadání
- platnost nabídky do
- oblast použití
- inventární číslo
- rok výroby
- elektrický příkon
- nositel nebo kontaktní osoba
- popis
- poznámky
- adresa URL (www)
- klíčová slova

Databáze byla volně přístupná na adrese <https://dnep.cvut.cz>.

Zadávání probíhalo offline. Záznamy mohli zadávat zaměstnanci ČVUT vyplněním formuláře, po revizi jedním z editorů byl záznam vystaven ve veřejné části[31].

Podářilo se získat ústní komentář pracovníka Výpočetního a informačního centra ČVUT v Praze, který je s databází obeznámen: *Databáze byla zpočátku úspěšná. Nepodařilo se však zavést efektivní procesy aktualizace údajů. Kvůli naakumulované neaktuálnosti obsahu byla databáze nakonec stažena z webu.*

I v současné době provozují jednotlivé univerzity katalogy vědeckého vybavení, obvykle pro svou interní potřebu. Příkladem může být Univerzita Jana Evangelisty Purkyně v Ústí nad Labem (UJEP) www.ujep.cz, kde je seznam vybavení poskytnut formou webové stránky.

5 Diskuze

Projekt Equipment.data.ac.uk byl zahájen s velkými ambicemi. Po předchozí „Equipment Database“ Univerzity v Loughborough a Kit-catalogue se jednalo o celostátní projekt. Záměrem bylo dle tvůrců vytvoření “One stop shop” (vše pod jednou střechou) pro sdílení vědeckého vybavení. Inovativní přístup zahrnoval automatizované zjišťování zdrojů dat (autodiscovery) s využitím linked data a automatizovaný sběr dat (harvesting).

V roce 2017 měl projekt již 51 členů a nabízel více než 13 tisíc položek vybavení. V následujících letech počet členů i položek narůstal, jak je uvedeno ve [Výsledcích analýzy dat](#).

Mezi roky 2022 a 2023 však došlo k podstatnému omezení vstupních a výstupních formátů, jak bylo uvedeno v oddílu [Historie projektu](#) a také k opuštění propojených dat (OPD) pro proces autodiscovery. Nově se organizace přihlašují prostřednictvím online formuláře a stejnou cestou hlásí organizace také změny v informacích o organizaci a umístění dat pro harvesting.

Nejpozději 13. 7. 2023 došlo k selhání funkce pro export agregovaných dat, která (navzdory kontaktování podpory Jisc) nebyla k datu odevzdání této práce zprovozněna.

Aktuálně tedy projekt vykazuje známky úpadku. Omezení vstupních a výstupních formátů lze považovat za projev racionalizace a sjednocení, ale pět měsíců nefunkční export agregovaných dat katalogu naznačuje, že další podpora projektu není prioritou provozovatele. Je otázkou, zda je příčinou celostátní rozsah (a tedy velký objem dat a zejména velký počet členských organizací a z toho plynoucí obtížnost koordinace). Mohlo by se zdát, že si provozovatelé zkrátka vzali “příliš velké sousto”. Na druhou stranu výše zmíněné projevy úpadku s počtem záznamů ani členských organizací na první pohled nesouvisejí - vyřazení funkcionalit a neschopnost opravy chyb ukazuje obecně spíše na podcenění podpory a vývoje, což může mít příčinu ve financování či obecně nízké podpory (v nefinančním smyslu) ze strany státu.

Samostatnou otázkou je důvod pro opuštění autodiscovery procesu, který reprezentoval inovativnost projektu. Důvodem mohou být nenaplněná očekávání (změny zásadních informací o institucích se nedějí tak často, aby si to žádalo automatizovaný proces), a v každém případě i úpravu OPD profilu musí někdo realizovat. Ovšem zatímco webový formulář může vyplnit i administrativní pracovník, pro úpravu RDF souboru vystaveného ve struktuře webových stránek instituce je již potřeba kvalifikovaný personál.

Také kvalita dat poskytovaných členskými organizacemi je sporná. Analýza dat v praktické části práce prokázala případy nedodržování polí povinných dle standardu UNIQIP, a to jak nevyplnění povinných polí, tak chybné vyplnění polí, kde jejich hodnota neodpovídá stanovenému účelu. Kontrolní mechanismy provozovatele dokáží část chyb odhalit (nevyplnění skupiny povinných údajů včetně identifikace konkrétních záznamů s chybou), otázka je, jak zjištěné chyby provozovatel s institucemi řeší. Předpoklad, že si odpovědní pracovníci instituce průběžně kontrolují funkcí “data check”, zjevně nepostačuje.

Navzdory zmiňovanému úpadku se katalog dále plní záznamy (průběh uveden v pododdílu [Databáze, počet záznamů a průběh v čase](#)), a také v posledních šesti měsících vstoupily do projektu tři nové organizace. Důvodem může být i to, že organizace čerpající grant UKRI jsou povinné při nákupu přístroje s cenou nad 138000 GBP nahlásit přístroje do katalogu. Celkově však z hlediska běžného provozu (zadávání nového vybavení a vyhledávání v katalogu) nic nebrání pokračování projektu. Nelze již patrně očekávat zásadní úpravy nebo vylepšení katalogu, ale pokud nenastanou okolnosti vyžadující velký zásah do kódu katalogu, může katalog dále poskytovat své služby a být přínosem pro sdílení vědeckého vybavení ve Spojeném království.

Obdobný projekt by byl jistě přínosem i pro Českou republiku. Přínosy sdílení vědeckého vybavení již byly zmíněny v oddílu [Přínosy](#):

Přínosy ekonomické:

- vyšší utilizaci, tedy efektivitu vynaložených prostředků

- úsporu z principu pořízení menšího počtu zařízení nebo možnost za stejné prostředky pořídit širší škálu přístrojů
- dosažení lepších cen a výhodnějších podmínek u dodavatelů

Přínosy ve veřejném zájmu, ve vědecké rovině

- zpřístupnění špičkového vědeckého vybavení i těm, kteří by jinak neměli možnost jej využít
- díky prostorové koncentraci zlepšení mezioborové spolupráce, předávání znalostí, což prokazatelně zvyšuje efektivitu výzkumné činnosti
- dostupnost technického personálu (v případě core utility) přispívá k edukaci uživatelů ohledně přístrojového vybavení, případně jim přímo zajišťuje podporu
- uživatelé působící v rámci core facility jsou díky kontaktu s víceoborovým prostředím a díky dostupnosti technického personálu lépe připraveni na další výzkumné aktivity i na budoucí profesní uplatnění

Do jisté míry již takové sdílení v ČR probíhá, a to zejména formou zařízení typu core facility na českých univerzitách a v dalších akademických institucích, včetně kooperativních projektů. Rozšíření na celostátní rozsah by nepochybně navýšilo pozitivní efekty sdílení, takový projekt by se ovšem potýkal se stejnými výzvami, jako britský Equipment.data.ac.uk. Zejména tedy podpora a financování ze strany státu a to nejen ve fázi budování a v období udržitelnosti, ale i dlouhodobě. Z dlouhodobého hlediska je nutná nejen údržba, ale i rozvoj systému.

Platformou, na které by taková služba mohla být provozována, je například IS VaVaI⁴⁹.

V současné době obsahuje IS VaVaI databáze CEA (aktivity VaVaI), VES (veřejné soutěže), CEP (projekty VaVaI) a RIV (výsledky VaVaI). Dále k poskytuje uživatelská rozhraní (pro zadávání dat) VaVER (pro příjemce podpory) a RoP (pro poskytovatele podpory). Data jsou veřejně dostupná, lze je exportovat i importovat ve standardizovaných formátech a je k nim

⁴⁹ Informační systém výzkumu, vývoje a inovací (IS VaVaI), <https://www.isvavai.cz/>, provozovaný Úřadem vlády ČR, shromažďuje informace o výzkumu, vývoji a inovacích podporovaných z veřejných rozpočtů v České republice a je jediným autorizovaným, úplným a závazným zdrojem těchto informací.

dispozici i strojový přístup přes API⁵⁰. [<https://www.isvavai.cz/>]

V rámci IS VaVaI by by bylo možné využít stávající databázi CEA, v rámci níž již existuje seznam velkých výzkumných infrastruktur, případně by vznikla nová databáze a rozhraní pro zadávání vybavení (pro členské organizace). Tak jako lze v současné době zadávat data do stávajících databází IS VaVaI importem XML dat, nepochybně by bylo možné takovou funkci vytvořit i pro databázi vybavení. Výhodou by bylo využití stávající technologie, kdy uživatelské rozhraní je za dobu provozu IS VaVai vyladěné a dle mého názoru z hlediska uživatelské přívětivosti patří ke špičce v rámci státní správy.

Pro toto řešení hovoří i fakt, že výzkumné organizace v ČR již mají se zadáváním do IS VaVaI zkušenosti. Z druhé strany vývojáři IS VaVaI již mají dostatečné zkušenosti pro implementaci další databáze a rozhraní.

Další výhodou tohoto řešení je fakt, že v rámci IS VaVaI již existuje registr výzkumných organizací, ten by tedy sloužil jako OPD bez nutnosti řešit profily organizací jako novou databázi.

Mohl by být projekt Equipment.data.ac.uk inspirací pro implementaci celostátního sdílení vědeckého vybavení v ČR? Patrně není důvod přebírat koncept autodiscovery a OPD, vzhledem k výše uvedené databázi organizací v rámci IS VaVaI. Stejně tak by patrně nebyla využita metoda harvesting. Pokud v současnosti výzkumné organizace zadávají jen do databáze RIV přes třicet tisíc výzkumných výsledků pomocí externích aplikací a API, menší příjemci pak pouze importem XML souborů nebo ručně, není patrně důvod, proč by stejně nemohli postupovat v případě výrazně nižšího počtu přístrojů ke sdílení.

Inspirativním základem pro koncept standardu (struktury dat) pro databázi vybavení by mohl být formát UNIQIP. Bude nutné zvážit rozšíření údajů přinejmenším o výrobce a model

⁵⁰ API, application programming interface, standardizované rozhraní pro poskytování dat nebo služeb vně informačního systému.

(jako je tomu v databázi Kit-Catalogue). Stejně tak by měla každá položka nést jedinečný identifikátor jako povinný údaj, stejně jako tomu je v jiných databázích IS VaVaI. Rozhodně lze doporučit použití standardizované kategorizace vybavení, což také odpovídá konceptu který je v IS VaVaI využíván (např. v databázi RIV je použita kategorizace dle FORD⁵¹).

V současné době je připravován *Návrh zákona o výzkumu, vývoji, inovacích a transferu znalostí a o změně některých souvisejících zákonů*⁵², který bude nově zakotvovat IS VaVaI. Ten zdá se směrem evidence vybavení zatím nemíří.

Otázkou je, zda je národní řešení potřeba v době, kdy existuje řešení celoevropského rozsahu (EOSC Marketplace). Namísto úsilí věnovanému tvorbě vlastního řešení by pak stát mohl vložit prostředky a lidské zdroje do řádného využití EOSC Marketplace s jednoznačným přínosem v celoevropském dosahu nabízeného vybavení a služeb.

⁵¹ FORD, Fields or Research and Development, je číselník zavedený OECD (Organizace pro hospodářskou spolupráci a rozvoj), <https://www.oecd.org/>.

⁵² <https://odok.cz/portal/veklep/material/ALBSCWXEC3V3/>

6 Závěr

Cílem této práce bylo upozornit na fenomén sdílení vědeckého vybavení a přispět k diskusi o možnostech realizace v České Republice.

Sdílení vědeckého vybavení přispívá k efektivnímu využití vědeckého vybavení, ale také podporuje mezioborovou spolupráci a zlepšuje dostupnost špičkových přístrojů i pro začínající vědce. Primárním nositelem sdílení jsou jednak tzv. core facility, tedy zařízení provozované akademickými institucemi, kde je koncentrováno různorodé vědecké vybavení určené pro využití jak v rámci instituce, tak pro externí subjekty. Druhou variantou sdílení vědeckého vybavení jsou konsorcia různého rozsahu, od regionálních až po celostátní.

Předpokladem pro realizaci sdíleného vybavení je v první řadě vhodné vybavení a prostory. Podstatný je také personál, který musí být schopen zajistit podporu uživatelům ať již formou edukace, nebo přímo asistence při práci s vědeckým vybavením. Pro sdílení vědeckého vybavení je také nezbytné, aby byla nabídka poskytovaného vybavení veřejně přístupná. Použité technologie se odvíjí od rozsahu poskytovaného vybavení a služeb, od prostého vystavení seznamu formou tabulky na webových stránkách instituce až po sofistikované informační systémy.

Významnou otázkou je financování, a to jak v případě konsorciálního sdílení, tak v případě core facility. Samotný provoz může být financován z plateb za využití vybavení, tento příjem však nemusí být dostatečný. Stát ovšem nehraje roli pouze ve financování, také by měl vytvářet příznivé legislativní podmínky pro sdílení vědeckého vybavení, které přispívá k efektivnímu využití prostředků, které stát vkládá do výzkumu a vývoje.

Důležitým faktorem je také motivace obou stran sdílení. Poskytovatel vybavení musí mít z této aktivity přínos, ať již čistě ekonomický z plateb za užívání přístrojů, nebo nepřímý. Motivace uživatele k využívání sdíleného vědeckého vybavení je různorodá. Zjevným přínosem je přístupnost špičkových technologií, které by jinak nemusely být pro uživatele

dostupné. Nejde však jen o přítomnost vybavení. V dobře uspořádaném zařízení jsou k dispozici odborníci na přístrojové vybavení, kteří uživatele proškolí, případně mu zajistí podporu během výzkumné aktivity či přímo daný experiment realizují.

Mimo zmíněných přínosů musí provozovatel zvážit také náklady a rizika, která sdílení vědeckého vybavení přináší. Je nutné zhodnotit ekonomickou bilanci, vzít v potaz vyšší opotřebení přístrojů a stanovit jasná pravidla pro využívání vybavení.

V rámci praktické části práce je formou případové studie popsán projekt otevřeného katalogu Equipment.data.ac.uk. Jsou popsány použité datové formáty a procesy, role propojených dat a následně byl analyzován obsah dat katalogu. Bylo zjištěno, že členské organizace vystavují ke sklizení i data, která nevyhovují standardu UNIQUIP. Kontrolní mechanismy sice část těchto chyb dokáží odhalit, otázkou však je, zda má provozovatel kompetence a vůli zjištěné chyby s členskými organizacemi řešit. Dále jsou posouzeny silné a slabé stránky projektu, je konstatován úpadek projektu.

Následuje přehled obdobných projektů v zahraničí i v České republice a jejich srovnání. Z tohoto srovnání vychází jako nejvíce pokročilý projekt ZEUS provozovaný vládou Korejské republiky.

V diskuzi je věnován prostor úvahám o možných příčinách úpadku projektu Equipment.data.ac.uk. a jsou nastíněny možné varianty dalšího vývoje. Navazuje zhodnocení možnost implementace v ČR s vizí realizace v rámci informačního systému IS VaVaI.

7 Soupis použité literatury a dalších zdrojů

- [1] ČESKÝ STATISTICKÝ ÚŘAD. *Přímá veřejná podpora výzkumu a vývoje v české republice: 2021*. Online. Praha, 2022. ISBN 978-80-250-3294-7. Dostupné z: <https://www.czso.cz/documents/10180/164606720/21100122.pdf/fe326bfb-15ae-4194-9f09-c31e327fcc98?version=1.1>. [cit. 2023-12-17].
- [2] STROM, T. Amanda; HAUGSTAD, Greg; SHU, Jonathan a SESHADRI, Ram. Shared instrumentation facilities: Benefiting researchers and universities, and sustaining research excellence. Online. *MRS Bulletin*. 2020, roč. 45, č. 5, s. 331-335. ISSN 0883-7694. Dostupné z: <https://doi.org/10.1557/mrs.2020.130>. [cit. 2023-11-22].
- [3] PARK, Eunwoo; LIM, Jaehyun; PARK, Byung Cheol a KIM, Daekeun. IoT-Based Research Equipment Sharing System for Remotely Controlled Two-Photon Laser Scanning Microscopy. Online. *Sensors*. 2021, roč. 21, č. 4. ISSN 1424-8220. Dostupné z: <https://doi.org/10.3390/s21041533>. [cit. 2023-11-22].
- [4] ANDEREGGEN, Stefan. Sharing Research Equipment to Bridge Intraorganizational Boundaries. *Research Technology Management* [online]. 2013, 56(1), 49-57. ISSN 0895-6308. Dostupné z: <https://doi.org/10.5437/08956308X5601082>. [cit. 2020-01-13],
- [5] GOULD, Julie. Core facilities: Shared support. Online. *Nature*. 2015, roč. 519, č. 7544, s. 495-496. ISSN 0028-0836. Dostupné z: <https://doi.org/10.1038/nj7544-495a>. [cit. 2023-11-26].
- [6] CHANG, Michael a GRIEDER, Franziska B. Sharing Core Facilities and Research Resources—An Investment in Accelerating Scientific Discoveries. Online. *Journal of Biomolecular Techniques: JBT*. 2016, roč. 27, č. 1, s. 2-3. ISSN 1524-0215. Dostupné z: <https://doi.org/10.7171/jbt.16-2701-004>. [cit. 2023-11-28].

- [7] HULSINK, Willem. Dedicated funding for leasing and sharing research and test facilities and its impact on innovation, follow-on financing and growth of biotech start-ups: the Mibiton case. *Venture capital* [online]. 2017, 19(1-2), 95-118. ISSN 1369-1066. Dostupné z: <https://doi.org/10.1080/13691066.2017.1261454>. [cit. 2020-01-13].
- [8] HONG, Jaekeun a CHUNG, Sunyang. User-oriented service and policy innovation in shared research equipment infrastructure: an application of the QFD and Kano's model to the Gyeonggi Bio-Center. Online. *Asian Journal of Technology Innovation*. 2013, roč. 21, č. 1, s. 86-107. ISSN 1976-1597. Dostupné z: <https://doi.org/10.1080/19761597.2013.810948>. [cit. 2023-11-23].
- [9] YOON, Donghun. The effect analysis of the research results on the spatial concentration and utilization sharing of research equipment. *International Journal of Engineering Business Management* [online]. 2017, 9, 184797901771035. ISSN 1847-9790. Dostupné z: <https://doi.org/10.1177/1847979017710350>. [cit. 2020-01-13].
- [10] JEON, Sehyeok; HWANG, Joo Sung; KIM, Miri a KIM, Seoyong. Analysis on the Implementation Plan to Increase the Joint Utilization Rate of National Research Facilities and Equipment. Online. In: 2022 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, 2022, s. 339-344. ISBN 978-1-6654-2197-3. Dostupné z: <https://doi.org/10.1109/BigComp54360.2022.00076>. [cit. 2023-11-22].
- [11] YOON, Donghun. The Research Ethics Policy for the Effective Utilization of Research Equipment. *International journal of technoethics* [online]. 2019, 10(2), 71-92. ISSN 1947-3451. Dostupné z: <https://doi.org/10.4018/IJT.2019070105>. [cit. 2020-01-13]
- [12] LEE, Kyoungmi a YANG, Jae-Suk. R&D cooperation in collaborative consumption of research equipment: An ERGM approach. Online. *Social Networks*. 2023, roč. 74,

- s. 118-126. ISSN 03788733. Dostupné z:
<https://doi.org/10.1016/j.socnet.2023.03.001>. [cit. 2023-11-22].
- [13] COX, A. a Ch. Gutteridge. equipment.data – Delivering a data autodiscovery infrastructure. CRIS2016: 13th International Conference on Current Research Information Systems (St Andrews, June 9-11, 2016). Dostupné z
<http://hdl.handle.net/11366/50>. [cit. 2019-05-24].
- [14] Data Autodiscovery—The Role of the OPD. *ISPRS International Journal of Geo-Information* [online]. 2016, 5(10), 167. ISSN 2220-9964. Dostupné z
<https://www.mdpi.com/2220-9964/5/10/167>. [cit. 2020-01-14].
- [15] HAUSENBLAS, Michael. 5 [hvězdička] OTEVŘENÁ DATA. Online. 2012.
Dostupné z: <https://5stardata.info/cs/>. [cit. 2023-12-17].
- [16] WOOD, David; ZAIDMAN, Marsha; RUTH, Luke a HAUSENBLAS, Michael.
Linked data: structured data on the web. Shelter Island: Manning, [2014]. ISBN 978-1-617290-39-8
- [17] RDF WORKING GROUP. *RDF: Resource Description Framework (RDF)*. Online. W3C. 2014. Dostupné z: <https://www.w3.org/RDF/>. [cit. 2023-12-17].
- [18] JOINT INFORMATION SYSTEMS COMMITTEE [JISC]. *About*. Online. JISC.
Jisc Equipment Data Service. 2023. Dostupné z: <https://equipment.data.ac.uk/about>. [cit. 2023-12-17].
- [19] *UK Research and Innovation FEC grants*. Online. In: UK RESEARCH AND INNOVATION. Dostupné z: <https://www.ukri.org/wp-content/uploads/2020/10/UKRI-201020-full-economic-costing-grant-terms-and-conditions-March-2020.pdf>. [cit. 2023-12-17].
- [20] *Kit-Catalogue Project*. Online. In: LOUGHBOROUGH UNIVERSITY. 2012.

- Dostupné z: <https://www.kit-catalogue.com/downloads/kitcat-booklet.pdf>. [cit. 2023-12-17].
- [21] GUTTERIDGE, Christopher J. *Launch of the first HE equipment-sharing database*. Online. In: UNIVERSITY OF SOUTHAMPTON. School of Electronics and Computer Science. 2023. Dostupné z: <https://www.ecs.soton.ac.uk/news/4216>. [cit. 2023-12-17].
- [22] JOINT INFORMATION SYSTEMS COMMITTEE [JISC]. *FAQ*. Online. JISC. Jisc Equipment Data Service. 2023. Dostupné z: <https://equipment.data.ac.uk/faq>. [cit. 2023-12-17].
- [23] DATA.AC.UK. *Organisation Profile Documents*. Online. 2015. Dostupné z: <https://opd.data.ac.uk/>. [cit. 2023-12-17].
- [24] JOINT INFORMATION SYSTEMS COMMITTEE [JISC]. *Jisc Equipment Data Service*. Online. 2023. Dostupné z: <https://equipment.data.ac.uk>. [cit. 2023-12-17].
- [25] JOINT INFORMATION SYSTEMS COMMITTEE [JISC]. *Equipment Data Service Submission Form*. Online. JOINT INFORMATION SYSTEMS COMMITTEE [JISC]. Jisc Equipment Data Service. 2023. Dostupné z: <https://forms.office.com/e/SKAEHrjvW4>. [cit. 2023-12-17].
- [26] DUCA, Daniela. *May: update on equipment.data*. Online. In: JOINT INFORMATION SYSTEMS COMMITTEE. Jisc. 2023. Dostupné z: <https://equipment.jiscinvolve.org/wp/2017/05/08/may-update-equipment-data/>. [cit. 2023-12-17].
- [27] JOINT INFORMATION SYSTEMS COMMITTEE [JISC]. *UNIQUEIP*. Online. JISC. Jisc Equipment Data Service. 2023. Dostupné z: <https://equipment.data.ac.uk/uniqueip>. [cit. 2023-12-17].
- [28] GUTTERIDGE, Christopher. *Jisc Digifest 2023*. Online. In: SOUTHAMPTON

- WEB AND DATA INNOVATION TEAM. <https://blog.soton.ac.uk/>. 2023. Dostupné z: <https://blog.soton.ac.uk/webteam/2023/03/08/jisc-digifest-2023/>. [cit. 2023-12-17].
- [29] EURO-BIOIMAGING ERIC. *About Euro-BioImaging*. Online. EURO-BIOIMAGING ERIC. Euro-BioImaging. Dostupné z: <https://www.eurobioimaging.eu/about-us/about-eubi>. [cit. 2023-12-17].
- [30] MATERIALS RESEARCH FACILITIES NETWORK [MRFN]. *Materials Research Facilities Network*. Online. Dostupné z: <https://mrfn.org/>. [cit. 2023-12-17].
- [31] ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ. *Databáze nabídek, expertů a přístrojů ČVUT*. Online. THE INTERNET ARCHIVE. Wayback machine. 2009. Dostupné z: <https://web.archive.org/web/20090202164451/http://www.dnep.cvut.cz/default.asp>. [cit. 2023-12-17].
- [32] ÚŘAD VLÁDY ČESKÉ REPUBLIKY. *Informační systém výzkumu, vývoje a inovací*. Online. C2021–2023. Dostupné z: <https://www.isvavai.cz/>. [cit. 2023-12-17].