

Stav správy výzkumných dat v České republice 2025 – výzkum a jeho výsledky

Martin Vávra, Tomáš Čížek, Jan Klusáček

Český sociálněvědní datový archiv Sociologický
ústav Akademie věd ČR, v. v. i.

1. 4. 2026



Spolufinancováno
Evropskou unií



Základní informace

- Výzkum byl zaměřen na téma správy výzkumných dat, potřeby odborníků, kteří s nimi pracují, a dopad iniciativy European Open Science Cloud CZ (EOSC CZ).
 - Sběr dat probíhal od 21. 5. 2025 do 28. 6. 2025.
 - Populace byla definována jako suma výzkumných projektů, o kterých existuje záznam v informačním systému VaVal, a které byly ukončeny v roce 2024 nebo zahájeny v období od roku 2020 do ledna 2025.
 - Ke každému projektu byl osloven jeden respondent, kterým byla osoba uvedená v systému jako hlavní řešitel. Dotazník vyplnilo 29,5 % (1121/3796) z oslovených osob. (2025-12-12)
- Dále prezentované výsledky jsou výstupem výzkumného šetření: Stav správy výzkumných dat v České republice 2025 .
- *Výzkum vznikl jako součást projektu Open Science I – Národní repozitářová platforma pro výzkumná data (NRP), realizovaného v rámci iniciativy EOSC CZ (<https://www.eosc.cz/>) s podporou MŠMT a ve spolupráci se Sociologickým ústavem AV ČR (<https://www.soc.cas.cz/cz/>).*

Kontext, cíle a východiska výzkumu

Témata výzkumu

- Mapování současné praxe RDM
- Identifikace vnímaných bariér ve správě dat
- Překážky pro sdílení a opětovné využití dat
- Institucionální podpora pro využívání dat třetími stranami
- Potřeby výzkumníků v kontextu správy dat

- Osobní údaje čerpány pouze z veřejných zdrojů, přesto byl projekt předložen Komisi pro etiku výzkumu SOU AV ČR.
- Účast byla dobrovolná a dobrovolnost uvedena v oslovovacím e-mailu.
- Účastníci obdrželi všechny informace o výzkumu, své roli i ochraně osobních údajů.
 - S ohledem na prioritní snahu získat dostatečnou návratnost jsme respondentům slíbili, že nebudeme data analyzovat na úrovni konkrétních pracovišť



Metodologie a průběh výzkumu

Metoda sběru dat

- Dotazníkové šetření s využitím standardizovaného online dotazníku (CAWI)
- Dotazník měl českou a anglickou verzi.
- Plán výzkumu byl posouzen Komisí pro etiku výzkumu SOU AV ČR.
- Osloveno bylo 3829 řešitelů výzkumných projektů
 - Šlo o projekty, o kterých existuje záznam v informačním systému VaVal, a které byly ukončeny v roce 2024 nebo zahájeny v období od roku 2020 do ledna 2025.

Průběh výzkumu

- Sběr dat proběhl 21. 5. 2025 do 28. 6. 2025 prostřednictvím online dotazníku
- V průběhu byly zaslány dvě připomínky výzkumu těm, kdo dotazník dosud nevyplnili
- Respondenti měli možnost posílat dotazy a připomínky na mailovou adresu výzkumu
- Data spolu s dokumentací jsou dostupná na: <https://doi.org/10.14473/CSDA/OJTMJF>

Zhodnocení vzorku, reprezentativita

- Dotazník vyplnilo 29,5 % (1121/3796) z oslovených osob (tj. těch, kteří obdrželi email s dotazníkem).
- Průměrně dotazník vyplňovali respondenti 16 minut.

Tabulka: Rozpis hodnot vstupujících do výpočtu návratnosti

emailů odesláno		3829
nezpůsobilých	chybná emailová adresa	28
	nesprávný adresát (bez vztahu k projektu)	5
způsobilí, nevyplněný dotazník	odmítnutí	40
	bez reakce	2521
	otevřeli, ale nevyplnili	114
způsobilí, vyplnili dotazník	částečné vyplnění	100
	úplné vyplnění	1021

Reprezentativita - typy výzkumných organizací

Tabulka: Srovnání struktury podle typu institucí, na nichž byly řešeny projekty zastoupené v datovém souboru (DS) a v databázi všech projektů využité pro oslovení respondentů (DO), v %

	Podíl (DS)	Podíl (DO)
Vysoké školy	59,3	59,9
AV ČR	26,4	26,3
Výzkumné instituce veřejné a státní (v.v.i.; organizační složky státu, příspěvkové organizace)	13,5	12,0
Ostatní (ne-veřejné, ne-státní výzkumné organizace)	0,8	1,8
Celkem	100,0	100,0

n (DS) = 1121, N (DO) = 3829

Charakteristiky respondentů

Pohlaví a věk respondentů

Tabulka: Pohlaví respondentů

	Počet	Podíl (%)
Muž	660	65,0
Žena	286	28,2
Nechci sdělit	69	6,8
Celkem	1015	100,0

Cílem výzkumu bylo získat data reprezentující v první řadě výzkumné projekty a teprve sekundárně pak populaci výzkumníků. Tabulky ale ukazují, že i variabilita respondentů vzhledem k věku a pohlaví je poměrně vysoká.

Tabulka: Věk respondentů

	Počet	Podíl (%)
25–34 let	78	7,6
35–44 let	372	36,5
45–54 let	346	33,9
55–64 let	127	12,5
65 a více let	49	4,8
Nechci sdělit	48	4,7
Celkem	1020	100

Postavení respondentů v rámci týmu

Tabulka: Postavení respondenta v rámci týmu

Vedoucí výzkumný pracovník	84,7 %
Výzkumný pracovník (Ph.D. a vyšší)	22,7 %
Doktorand/ka	1,0 %
Student bakalářského, magisterského stupně	0,1 %
IT profesionál (administrátor, developer, support)	0,9 %
Datový specialista, datový steward	2,8 %
Konzultant, podpora, administrativa	2,7 %

N = 1121; Poznámka: respondent mohl uvést i více pozic, které v týmu zastává, proto součet % přesahuje 100.

- Oslovení mohli podle pokynů přeposlat dotazník k vyplnění jinému členu týmu
- Přesto výrazná většina respondentů jsou "principal investigators", tedy ti, které jsme oslovili. Naopak datových specialistů jsou v datech necelá 3 %

Charakteristiky výzkumných projektů

Typy výzkumných projektů

	Počet	Podíl (%)
Základní výzkum	661	59,0
Aplikovaný výzkum	302	26,9
Experimentální vývoj	96	8,6
Infrastruktura výzkumu, experimentálního vývoje a inovací	62	5,5
Celkem	1121	100,0

- I zde jsou podíly ve vzorku velmi blízké těm v základním souboru

Vědní obor v kterém tým respondenta působí

	Počet	Podíl (%)
Přírodní vědy	511	45,6
Technické vědy	174	15,5
Lékařské vědy	111	9,9
Zemědělské a veterinární vědy	58	5,2
Společenské vědy	115	10,3
Humanitní vědy	121	10,8
Jiný obor (volná odpověď)	31	2,8
	1121	100,0

- Založeno na deklaraci respondenta

Velikost projektových týmů popisovaných ve výzkumu

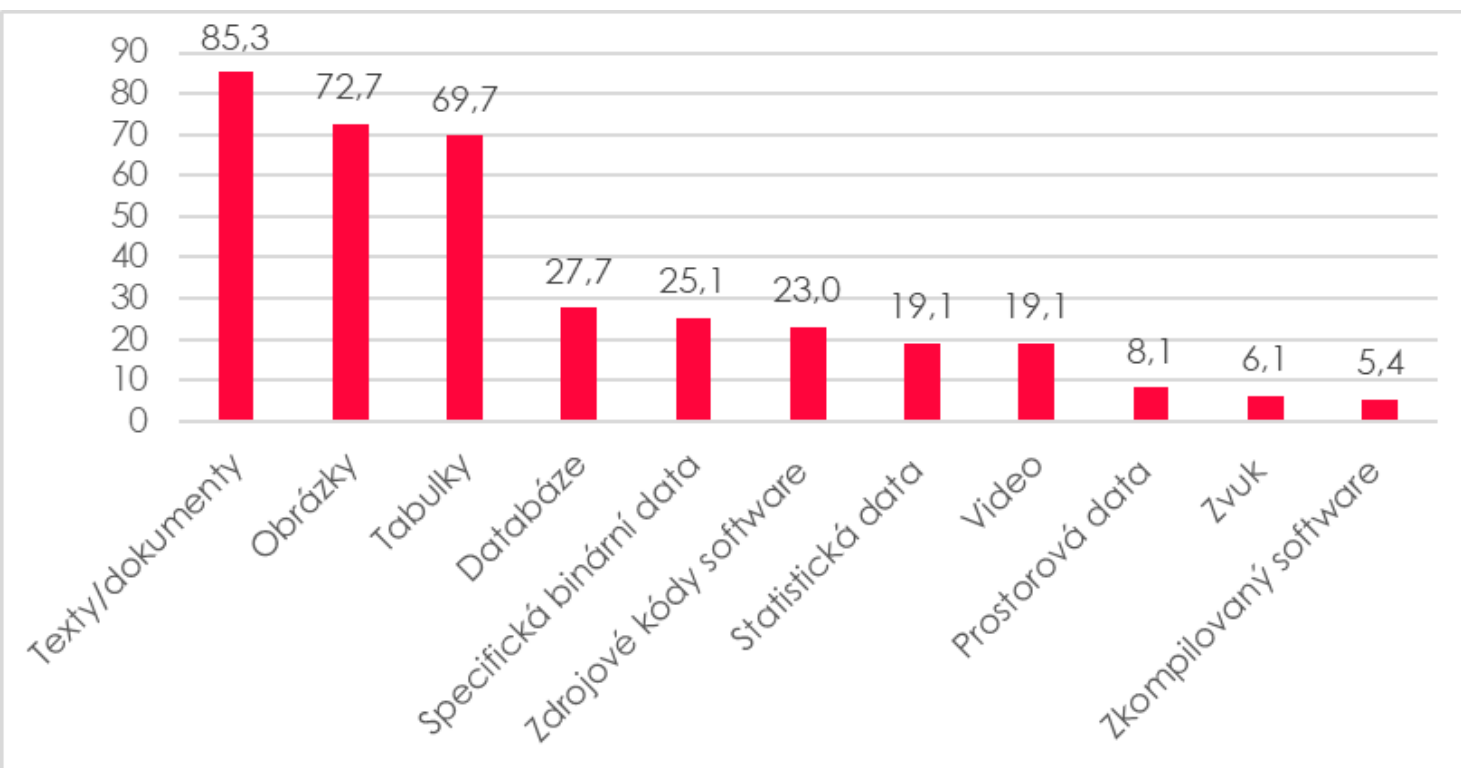
Tabulka: Velikost projektového týmu

	Počet	Podíl (%)
Méně než 5	298	26,6
5–10	570	50,8
11–20	187	16,7
21–30	30	2,7
Více než 30	36	3,2
Celkem	1121	100,0

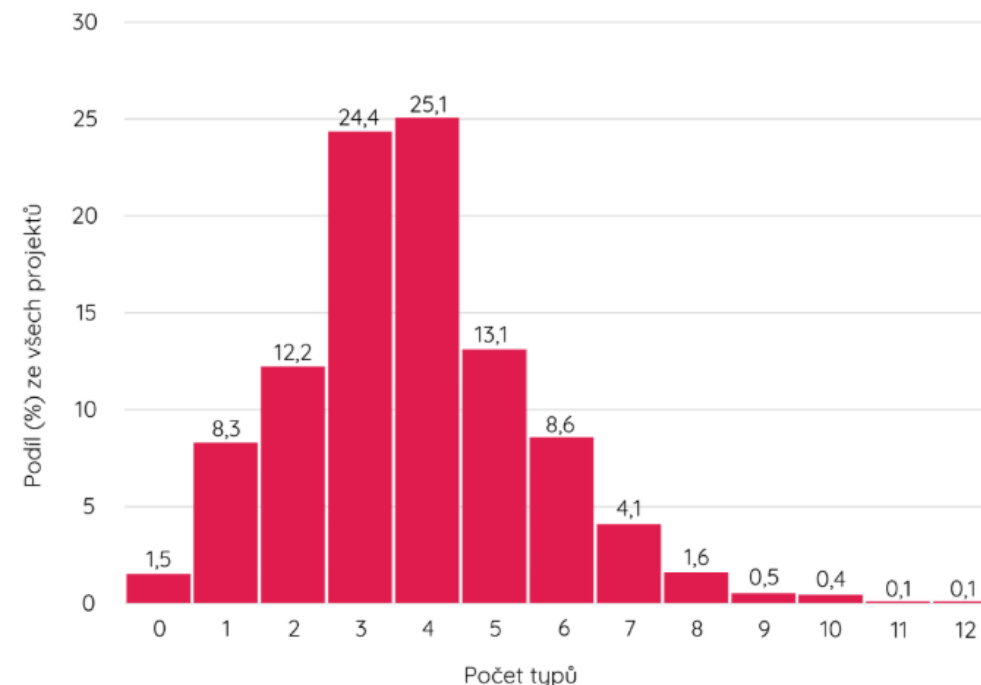
- Největší týmy spadají do oblasti infrastrukturních projektů, zatímco v průměru nejmenší jsou ty v základním výzkumu.

Generování výzkumných dat

Typy dat generované v rámci projektu



Graf: Počet typů dat produkovaných v projektech



Texty, obrázky a tabulky jako typy dat dominují. Tabulky, obrázky a texty také představují nejčastěji se vyskytující kombinaci datových typů produkovaných v rámci jednoho projektu.

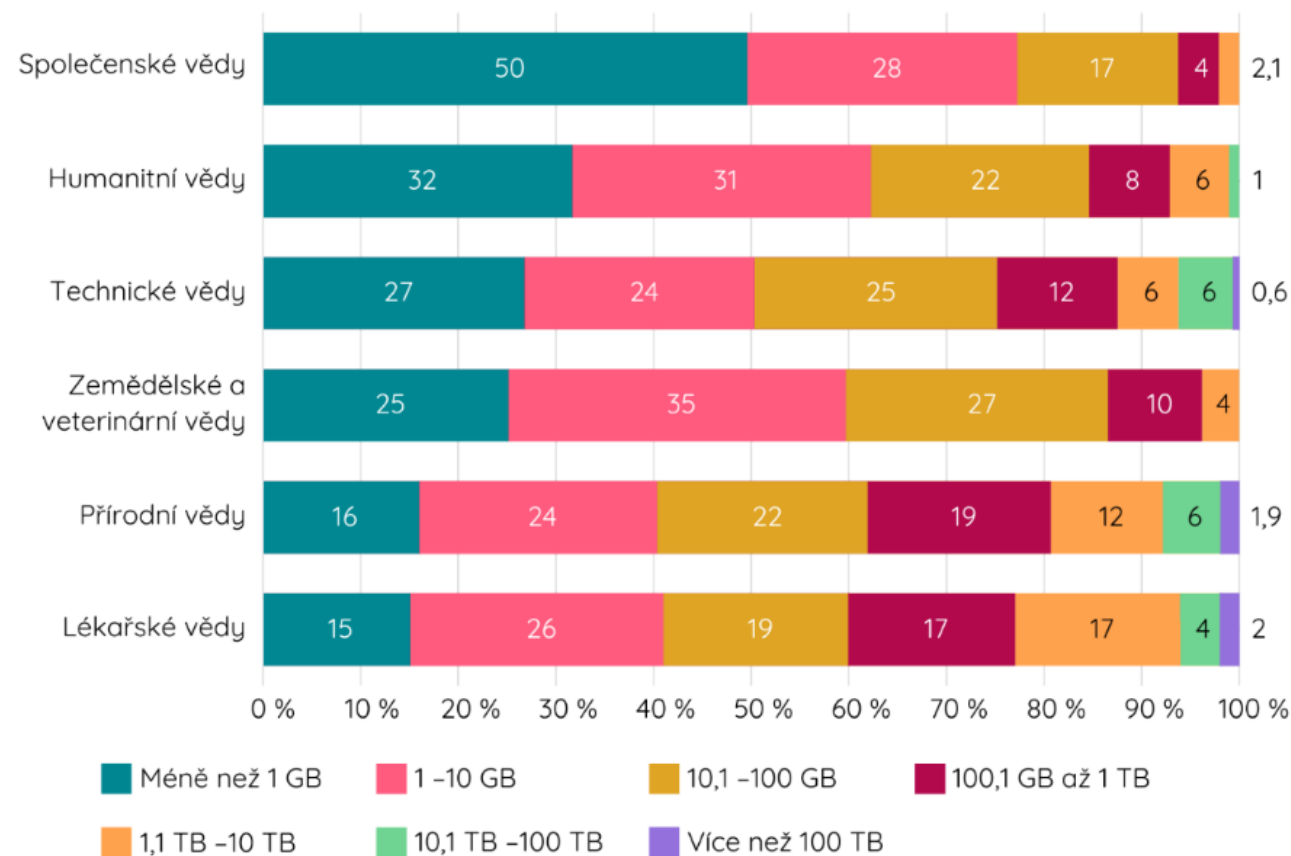
Některé typy dat, zejména texty, jsou produkovány ve všech oblastech vědy zatímco u ostatních jsou výrazné oborově specifické rozdíly, například v technických, přírodních a v menší míře i medicínských vědách jsou významně častěji produkována "softwarová" data.

Jen velmi malý podíl projektů neprodukuje žádná data (jsou to buď infrastrukturní projekty nebo projekty "teoretických" disciplín (např. filosofie, některé oblasti matematiky atp.)

Množství dat vytvořené v rámci projektu za rok

	Počet	Podíl (%)
Méně než 1 GB	233	21,1
1 – 10 GB	260	23,6
10,1 – 100 GB	219	19,8
100,1 GB až 1 TB	144	13,0
1,1 TB – 10 TB	93	8,4
10,1 TB – 100 TB	42	3,8
Více než 100 TB	12	1,1
Nevím	101	9,1
Celkem	1104	100,0

K celkovému množství produkováných dat v projektu nejvíce přispívají konfigurační data (INI, CONF...) a zkompilovaný software (binární spustitelné soubory), naopak nejméně statistická data ((DTA, POR, SAS, SAV...).



Citlivá data v projektech

Tabulka: S jakými typy citlivých dat v projektu pracují

	Ano
Osobní údaje (např. jméno, rodné číslo, zdravotní stav, biometrické údaje)	20,0 %
Obchodní tajemství (např. informace o klientech, strategické plány, finanční dokumenty)	6,0 %
Utajovaná data (např. data týkající se bezpečnosti státu)	1,2 %

N = 1121; Odpověď na otázku: „Pracujete ve Vašem projektu s těmito typy citlivých dat?“

S osobními daty pracuje jen menšina projektů,
oborové rozdíly jsou ovšem výrazné



Citlivá data v projektech

- Osobní data jsou nejčastěji zpracovávána v lékařských vědách
- Obchodní tajemství pak ve vědách technických

Tabulka: Podíl projektů produkujících citlivá data v závislosti na oboru vědy (čísla v buňkách indikují podíl projektů v dané oblasti vědy, které vytvářejí uvedené typy citlivých dat)

	Přírodní vědy	Technické vědy	Lékařské vědy	Zemědělské a veterinární vědy	Společenské vědy	Humanitní vědy
Osobní údaje (např. jméno, rodné číslo, zdravotní stav, biometrické údaje)	8,8	13,2	64,9	3,4	35,7	24
Obchodní tajemství (např. informace o klientech, strategické plány, finanční dokumenty)	3,9	17,8	1,8	6,9	5,2	0,8
Utajovaná data (např. data týkající se bezpečnosti státu)	0,0	3,4	0,9	0,0	2,6	1,7

Využívání metadat pro popis dat

- Ve výzkumu 59 % respondentů uvedlo , že v popisovaném projektu využívá metadata pro popis dat
- Mezi způsoby vytváření metadat dominuje zápis v readme souboru nebo podobném formátu

Tabulka: Využívané způsoby vytváření metadat v projektu

	Podíl projektů, které využívají
Používáme formulář pro vkládání metadat dle specifikovaných metadatových schémat (např. DublinCore, Data Cite, DCAT, DDI,...)	7,2 %
Zapisujeme důležité informace v readme (textový soubor) nebo podobném formátu	36,3 %
Metadata k datasetu jsou generována automaticky v přístroji	23,2 %

260

N = 1121; Odpověď na otázku: „Jaký je hlavní způsob vytváření metadat ve Vašem projektu?“

Důvody, proč nevytvářejí metadata

- Zde převládá odpověď, že pro daný projekt není vytváření metadat důležité

Tabulka: Důvody, proč nepoužívají metadata pro popis dat

	Respondent souhlasí
Nevíme, jak na to	6,2 %
Nemáme potřebné technické vybavení	2,4 %
Nemáme dostatečné kapacity (lidské, finanční, časové)	10,7 %
Pro projekt to není důležité	28,8 %

N = 1121; Odpověď na otázku: „Proč svá data nepopisujete?“

Shrnutí části "Generování výzkumných dat"



- Drtivá většina z těch, kteří odpověděli na dotazník, produkuje v rámci projektu, na jehož základě jsme je oslovili, nějaká výzkumná data (pouze 1,5 % uvedlo explicitně, že žádná taková data nevytvářejí).
- Některé z typů dat procházejí všemi oblastmi vědy (texty, obrázky, tabulky), využití jiných je více diferencované (strukturované databáze, statistická data, video), další jsou oborově poměrně specifické (software, prostorová data).
- Množství dat vyprodukované v sledovaných projektech za rok je u většiny projektů do 100 GB (64 % odpovědí). Projektů, které vyprodukují více než 100 TB dat za rok, je v našem vzorku jen 1 %.
- V našem souboru je 74 % respondentů, kteří uvedli, že nepracují s citlivými daty.
- V našem výzkumu 59 % respondentů uvedlo, že v projektech, na které jsme se ptali, využívají metadata k popisu dat. Nejčastější formou tohoto popisu je ovšem využívání textových souborů typu „readme“, spíše než využití metadatových standardů. Standardy používá 7 % respondentů. 23 % pak uvádí, že metadatové záznamy produkují přímo jejich výzkumné přístroje.

Uchovávání dat

Povinnost ukládat výzkumná data

- Vysoké procento respondentů považuje ukládání dat za povinnost danou z různých důvodů jako je například podmínka poskytovatele grantových prostředků, z důvodu publikace výsledků a jiných relevantních důvodů, nicméně 21 % respondentů se s takovým požadavkem nesetkalo.



Vědomí povinnosti ukládání dat vhledem k datu zahájení projektu

Tabulka: Povinnost uchovávat výzkumná data ve vztahu k roku zahájení projektu. V buňkách uvedeno % respondentů, kteří s výrokem souhlasí

	2020	2021	2022	2023	2024	2025
Ano, je to podmínka poskytovatele financování	33,3	35,6	36,7	42,9	57,4	48,6
Ano, je to podmínka vydavatele/časopisu	18,5	21,2	21,9	20,9	25,5	23,7
Ano, je to podmínka domovské instituce nebo organizace	25,9	23,7	31,0	22,3	23,0	31,6
Ano, je to součást plánu správy dat	25,9	35,6	30,0	40,4	52,8	50,3
Ano, na základě spolupráce se zahraničními partnery	14,8	11,0	7,1	7,9	7,7	9,6
Ne, nesetkali jsme se s tímto požadavkem	22,2	27,1	32,9	19,5	12,3	15,8

Ukládání dat v průběhu projektu

- Jenom polovina dotázaných používá nějaké formy institucionálního úložiště v průběhu řešení projektu, převažuje ukládání na osobní počítače (67 %) a více než polovina ukládá data na média náchylná k poškození nebo ztrátě jako jsou USB disky a podobné.

Tabulka: Úložiště dat v průběhu projektu

	Počet	Podíl (%)
USB, externí disk, CD/DVD, paměťová karta	596	53,2
Osobní počítač / notebook	744	66,4
Místní úložiště vaší instituce (např. NAS)	538	48,0
Institucionální cloud (...)	577	51,5
Osobní nebo free cloud (...)	240	21,4
Národní infrastruktura (např. e-INFRA CZ / CESNET)	162	14,5
Analogové dokumenty (např. papírové dokumenty)	176	15,7
Fyzické objekty (např. chemické vzorky)	178	15,9

Zálohování a verzování dat

Tabulka: Zálohování dat

	Počet	Podíl (%)
Ano, zálohujeme pravidelně	692	64,8
Ne, ale občas zálohujeme	333	31,2
Ne	43	4,0

Tabulka: Verzování dat

	Počet	Podíl (%)
Ano, vytváříme nové verze a původní data uchováváme	467	44,1
Ano, přepisujeme původní data při vytváření nových verzí	57	5,4
Ne, data vznikají jednorázově a dále se neupravují	429	40,5
Ne, přepisujeme původní data bez verzování	106	10,0

- Více než polovina dotázaných ukládá svá data do repozitářů. Třetina z těchto dotázaných pak volí mezinárodní univerzální repozitáře jako je například Zenodo.

Tabulka: Nahrávání dat do repozitáře

	Počet	Podíl (%)
Ano	578	54,1
Ne	296	27,7
Nevím	194	18,2

N= 1068

Tabulka: Volba repozitáře

	Počet	Podíl (%)
Oborový repozitář (např. ČSDA, PDB, NOMAD)	118	10,5
Univerzální repozitář na vaší instituci/univerzitě (např. DSPACE)	205	18,3
Univerzální repozitář mimo vaší instituci/univerzitu (např. Zenodo, Figshare)	370	33,0
V rámci tohoto projektu neukládáme/neplánujeme uložit data do repozitáře	37	3,3
Nevím, ještě jsem o tom nepřemýšlel/a	152	13,6

N= 770

Očekávaná doba uložení dat

- Většina dotázaných (téměř 60 %) plánuje svá data ukládat trvale.

Tabulka: Očekávaná doba uložení dat (%)

	Přírodní vědy	Technické vědy	Lékařské vědy	Zemědělské a veterinární vědy	Společenské vědy	Humanitní vědy
Do 3 let	1,6	7,7	2,8	10,9	4,5	5,3
3 – 5 let	14,3	22,6	22,2	21,8	20,9	10,6
6 – 10 let	20,5	23,2	21,3	21,8	13,6	5,3
Více než 10 let	18,9	11,9	28,7	12,7	10,0	9,7
Trvale	43,0	28,6	25,0	32,7	41,8	59,3
Nechceme / nepotřebujeme data uchovávat	1,6	6,0	0,0	0,0	9,1	9,7

Sdílení a opakované využití dat

Využití výzkumných dat jiných vědců nebo institucí

- Nejčastěji využívají data jiných vědců či institucí pro sekundární analýzu výzkumníci v oblasti přírodních a technických věd, oproti tomu nejméně často využívají tato data ti, kdo provádějí výzkum v zemědělských a veterinárních vědách.
- Pro oblast společenských věd jako pro jedinou platí, že nejčastěji používaným způsobem získání dat, je stažení z volně přístupného repozitáře.

Tabulka: Využil někdy respondent data z následujících zdrojů?

	Ano
Někdy jsem stáhl/a data z volně přístupného repozitáře nebo úložiště	58,3 %
Někdy jsem získal/a data od jiného vědce/vědkyně na základě individuální žádosti	64,3 %
Někdy jsem použil/a data od komerčních institucí nebo organizací	30,6 %

N = 1054

Využívání dat veřejné zprávy

Tabulka: Využil/a jste někdy pro svůj výzkum veřejná nebo otevřená data od státu nebo veřejné správy?

	Počet	Podíl (%)
Ano, více než jednou	290	27,6
Ano, jednou	32	3,0
Ne	640	60,9
Nevím	89	8,5
Celkem	1051	100,0

Tabulka: Máte zájem využívat data veřejné správy, která jsou chráněná speciálními předpisy (např. rejstříky, registry, dotazování podle zákona o svobodném přístupu k informacím)?

	Počet	Podíl (%)
Ano	251	23,9
Ne	519	49,4
Nevím	281	26,7
Celkem	1051	100,0



Sdílení vlastních dat s dalšími vědci

Tabulka: Jakým způsobem respondenti výzkumu sdíleli svá výzkumná data (v %)

	Ano, často	Ano, párkrát	Ano, jednou	Ne
Neformálně s jinými vědci v rámci ČR (např. emailem, osobním úložištěm)	36,4	47,2	1,4	15,1
Neformálně s jinými vědci ze zahraničních pracovišť (např. emailem, osobním úložištěm)	30,3	48,3	2,7	18,7
Prostřednictvím datového úložiště s otevřeným přístupem	19,0	23,4	3,8	53,8
Prostřednictvím datového úložiště s chráněným přístupem	19,3	30,5	3,4	46,8

- Neformální sdílení je častěji využívanou cestou, ovšem i datová úložiště pro sdílení využívá okolo poloviny respondentů, nejméně často jsou využívána datová úložiště s otevřeným přístupem
- Nejčastěji sdílejí všemi uvedenými způsoby svá data respondenti z oblasti přírodních věd, následování vědci v oblasti technických věd.

Cesty sdílení dat

Tabulka: Jakými způsoby respondenti svá data sdílejí

	Ano, využívá tuto cestu (%)
E-mailová korespondence	57,3
Institucionální cloudová úložiště (např. Dropbox, Google Drive, Microsoft OneDrive/SharePoint)	44,9
Osobní nebo free cloudová úložiště (např. GitHub, Dropbox, Google Drive, Microsoft OneDrive/SharePoint)	41,2
Místní úložiště vaší instituce (např. NAS)	26,1
Univerzální repozitář (např. Zenodo, Figshare)	22,3
Fyzické objekty (např. chemické vzorky)	20,4
Národní infrastruktura (např. e-INFRA CZ / CESNET)	13,5
Analogové dokumenty (např. papírové dokumenty v krabicích)	9,9
Oborový repozitář (např. ČSDA, LINDAT, PDB, NOMAD)	7,8

Využívání licencí k datům

Tabulka: Využívá respondent licence k datům?

	Počet	Podíl (%)
Ano, vím, jaké licence potřebuji	344	33,2
Ne, protože se v problematice neorientuji	284	27,4
Ne, není to pro mě relevantní	409	39,4
Celkem	1037	100,0

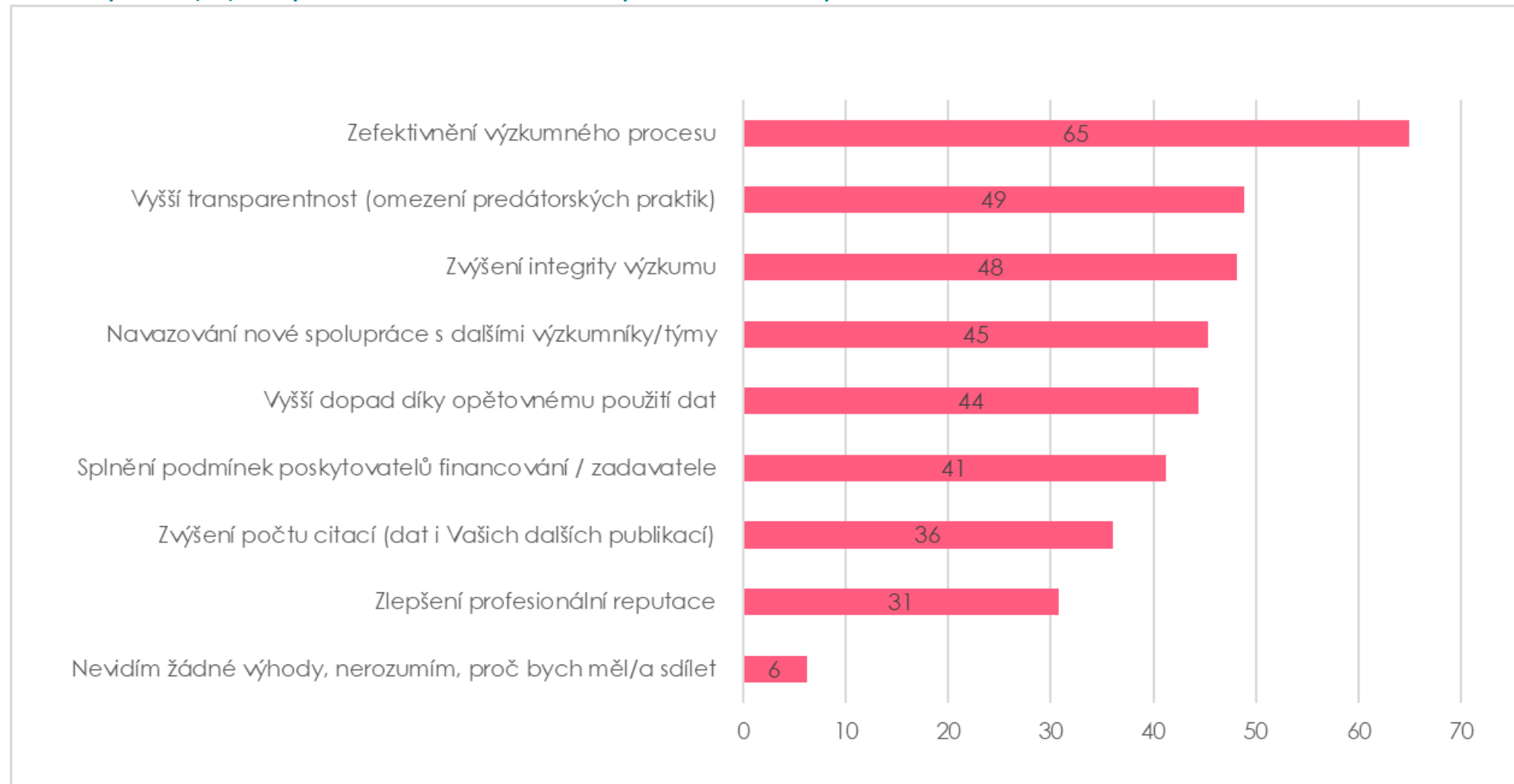
Tabulka: Jaké licence respondenti využívají

	Počet	Podíl (uvádí pro každou licenci podíl respondentů, kdo danou licenci využívají, %)
Creative Commons	316	28
MIT	42	4
Apache License	18	2
GNU General Public License	70	6
BSD	4	0

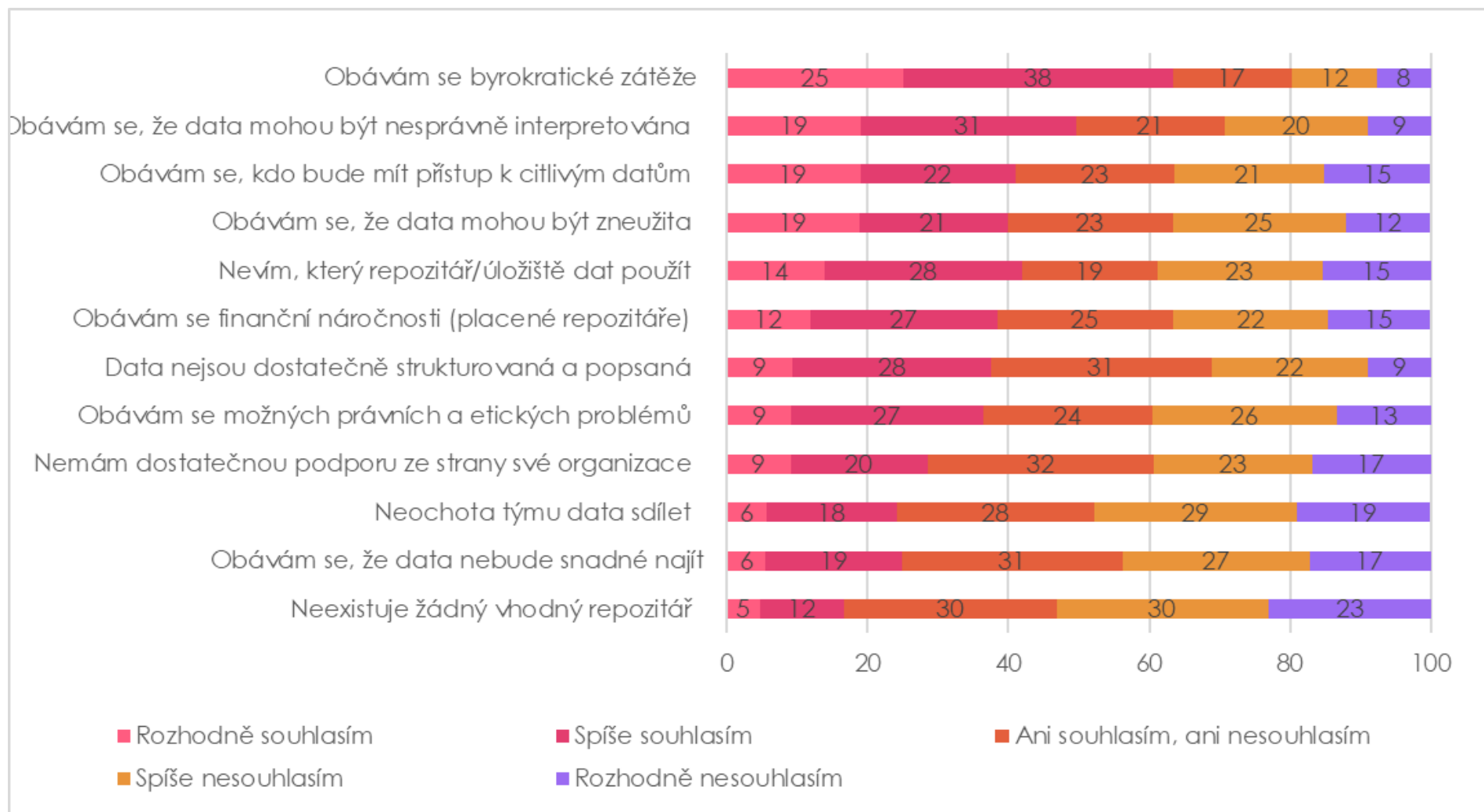
N = 1121

Vnímané benefity sdílení dat

Graf: podíl (%) respondentů, kteří zvolili příslušnou odpověď



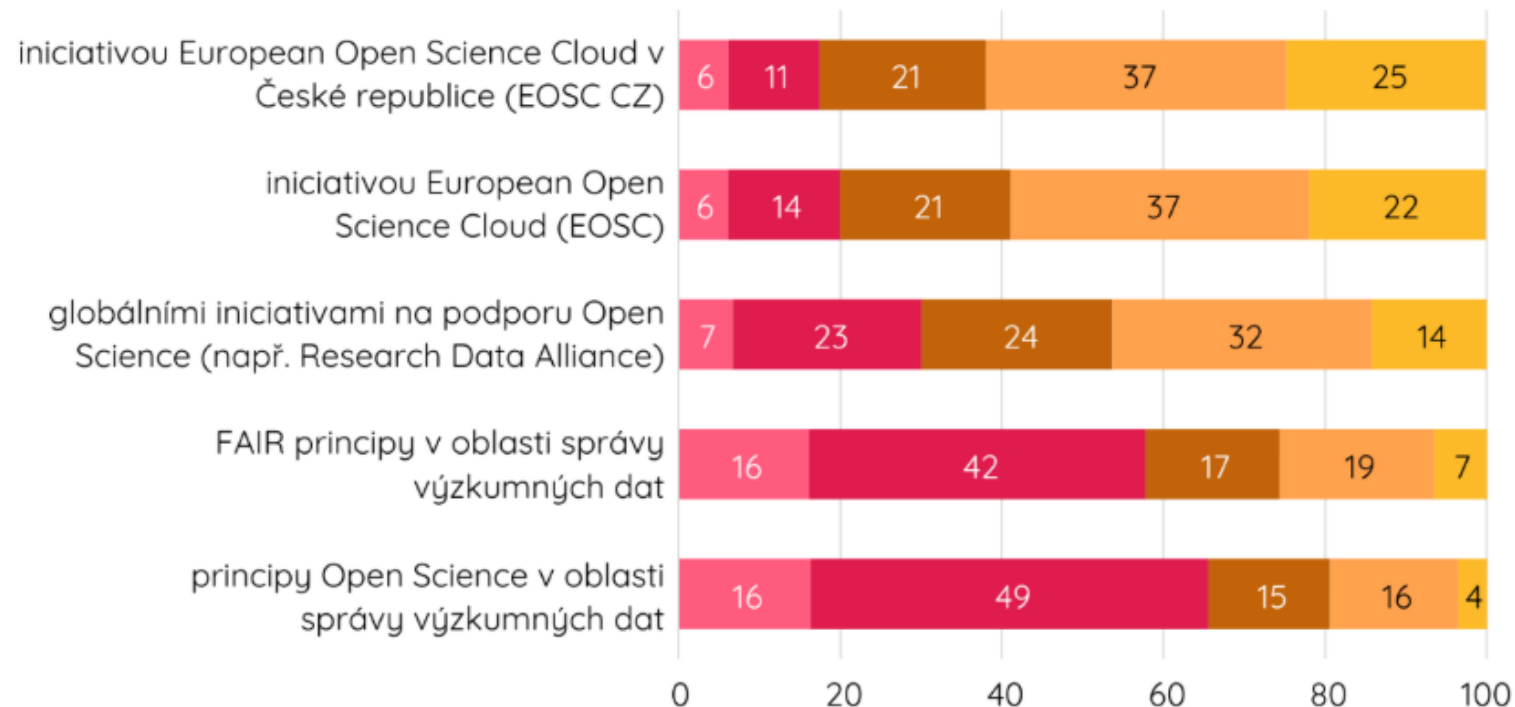
Vnímané komplikace sdílení dat



Znalost principů a institucí otevřené vědy

Graf: podíl (%) respondentů, kteří zvolili příslušnou odpověď

Obeznamenost s:



 Rozhodně souhlasím  Spíše souhlasím  Ani souhlasím, ani nesouhlasím
 Spíše nesouhlasím  Rozhodně nesouhlasím

- Znalost principů otevřené vědy je podstatně lepší než znalost jejích institucí
- Nejčastěji uvádějí znalost respondenti z oblasti technických a přírodních věd, nejméně často pak ti z oblasti věd zemědělských a veterinárních
- Dostupnost data stewarda na pracovišti respondenta pozitivně souvisí s vyšší obeznameností s principy a institucemi OS

Shrnutí části "Sdílení a opakované využití dat"



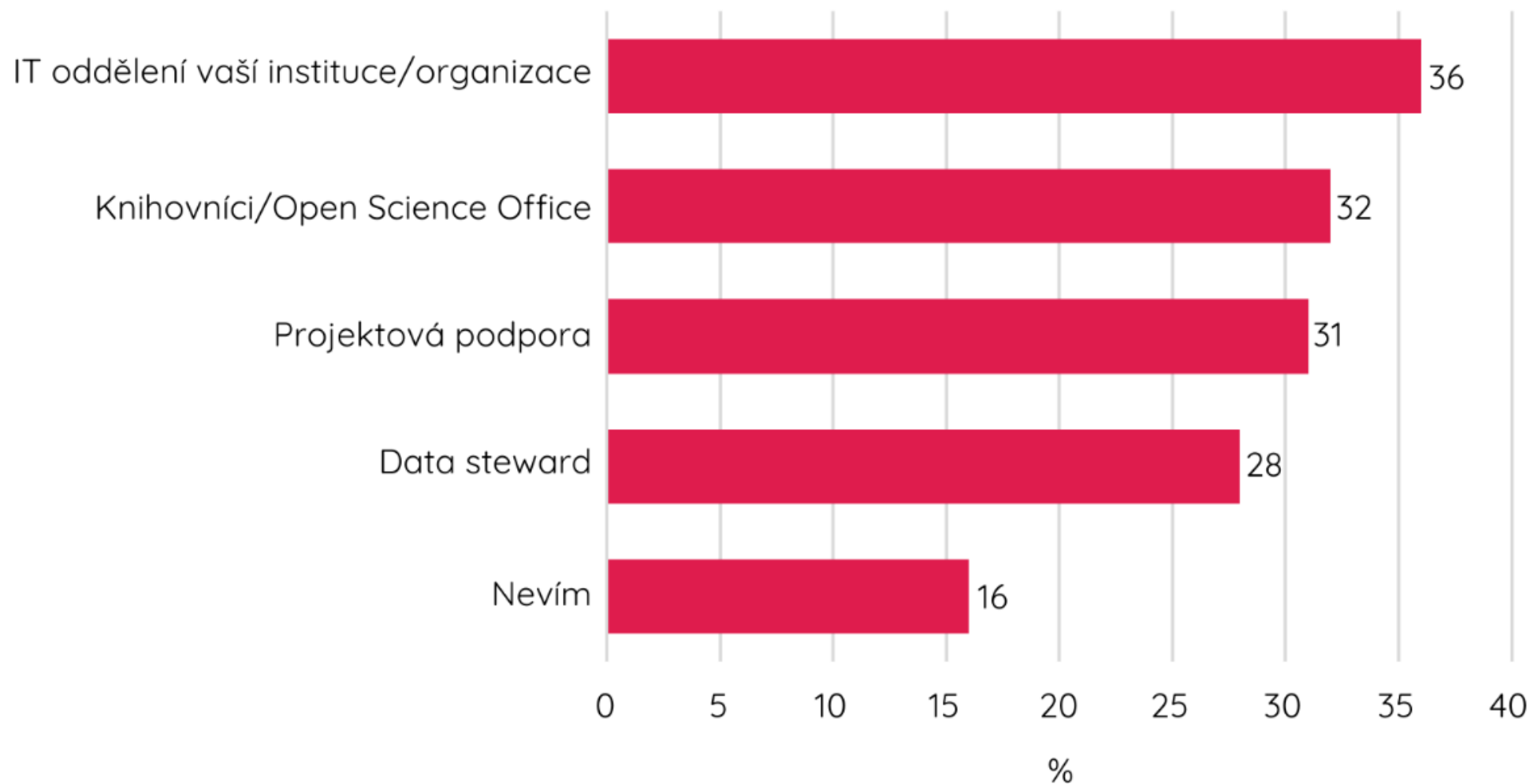
- Od jiných výzkumníků získávají respondenti data nejčastěji na základě individuální žádosti, významně častěji než prostřednictvím repozitářů. Nejčastěji využívají data jiných vědců či institucí pro sekundární analýzu výzkumníci v oblasti přírodních a technických věd.
- Také pro aktivní sdílení dat s ostatními vědci jsou využívány více neformální cesty (nejčastěji email) a individuální domluva, spíše než sdílení prostřednictvím repozitáře.
- Data veřejné správy využila o něco méně než třetina respondentů (v sociálních vědách ovšem 58 % výzkumníků).
- Licence pro data používá třetina respondentů. Dalších 27 % ovšem licence nevyužívá proto, že se v problematice neorientuje. Nejčastěji používanou licencí je s velkým odstupem Creative Commons.
- Mezi benefity sdílení dat je nejčastěji uváděno „zefektivnění výzkumného procesu“ (65 % respondentů uvedlo), nejméně často pak „zlepšení profesionální reputace“ (31 %).
- Jako překážka sdílení dat jsou v rámci celého souboru respondentů nejčastěji uváděny obavy z byrokratické zátěže
- Deklarovaná obeznámenost se základními principy otevřené vědy je podstatně lepší než znalost jejích institucí.

Podpora a rozvoj kapacit

Podpora a rozvoj kapacit

- institucionální podpora správy dat
- vytváření plánů správy dat
- potřeby výzkumníků
- postoje vědců k otevřené vědě

Koho mají k dispozici při správě dat?

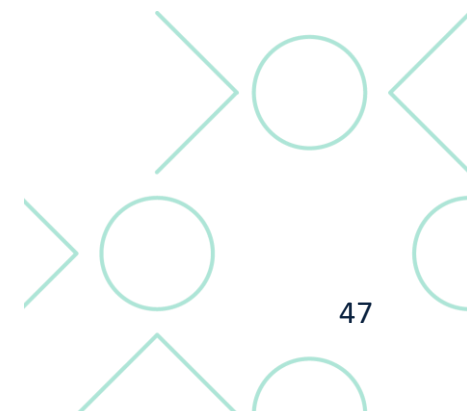
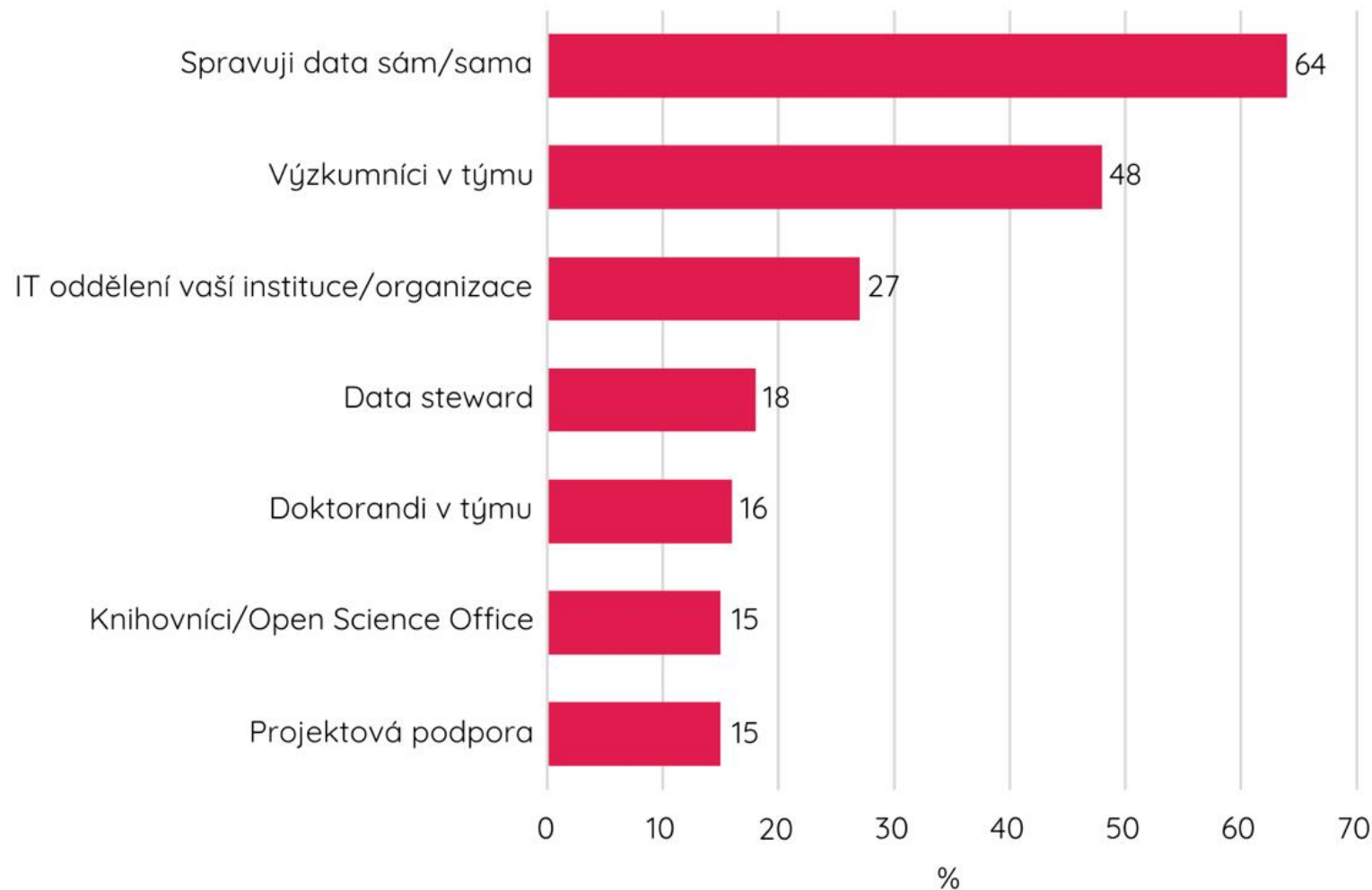


Podpora správy dat dle kategorie instituce

	VŠ (N=665)	AV ČR (N=296)	Veřejné, příspěvkové, státní (N=151)
Projektová podpora	34	26	28
Data steward	28	35	13
Knihovníci/Open Science Office	34	35	19
IT oddělení vaší instituce/organizace	31	39	48
Nevím	16	14	16

N = 1112, soukromé instituce/organizace nezahrnuty kvůli nízkému zastoupení

Kdo reálně spravuje data?



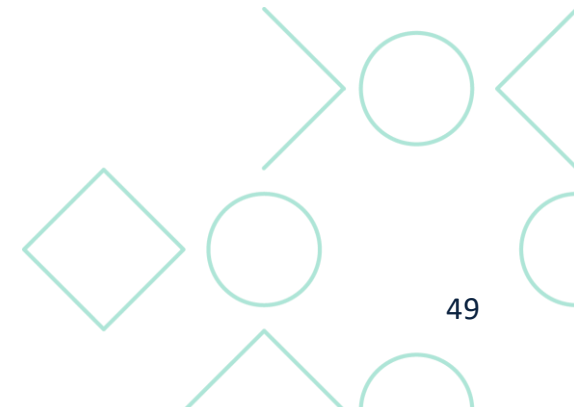
Kdo reálně spravuje data?

- AV ČR -> častější data stewardi
- veřejné, příspěvkové, státní instituce -> častější IT oddělení
- podle dostupnosti -> data steward a IT oddělení před knihovníky/Open Science Office a projektovou podporou
- spokojenost při správě dat podle spravujícího subjektu
 - nejspokojenější s doktorandy a kolegy
 - menší spokojenost s IT odděleními



Data management plan

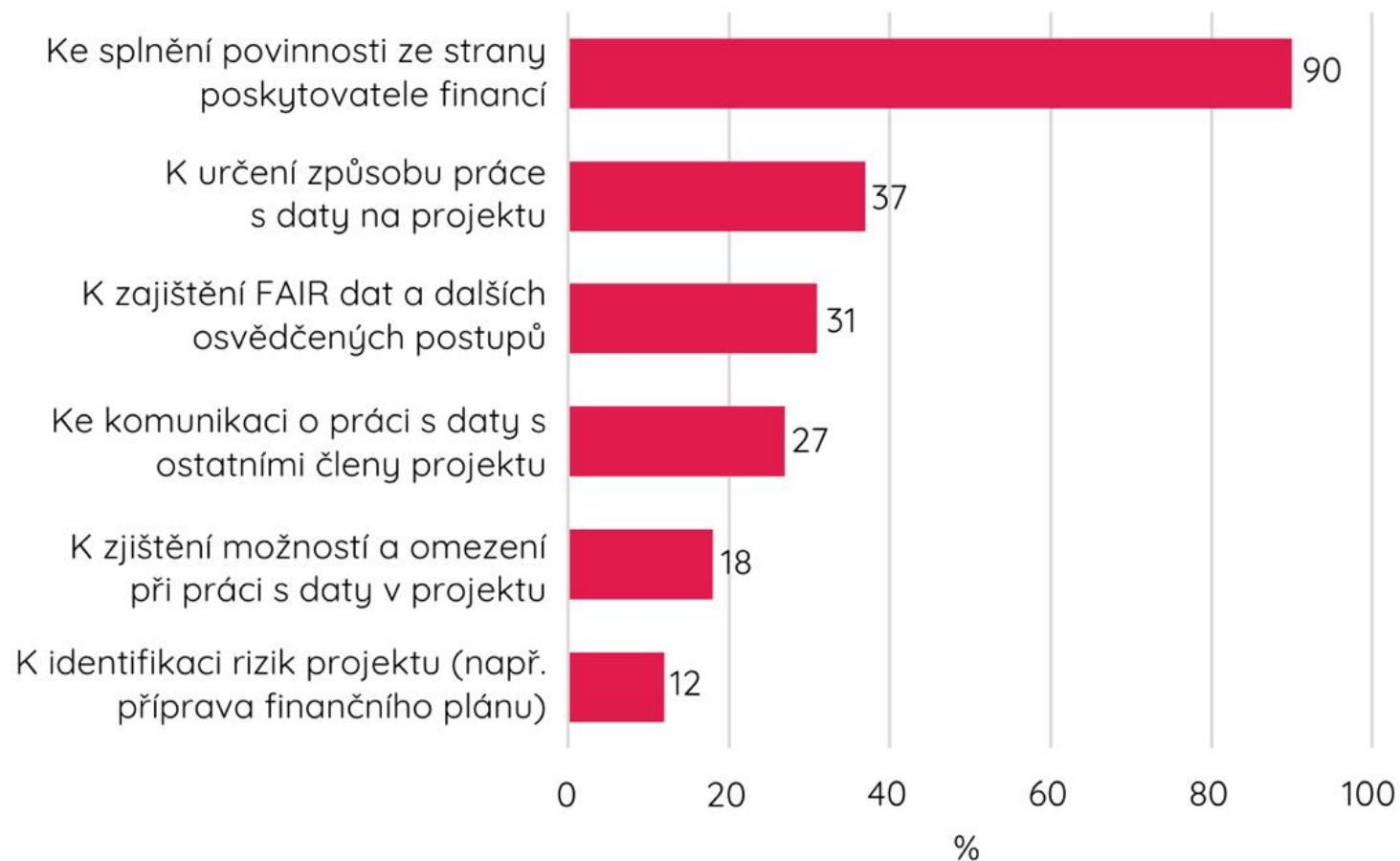
- 60 % respondentů vytvořilo nebo plánuje vytvořit
- převážně nové projekty (>70 %)
- podle instituce
 - ústavy AV ČR více
 - univerzity (technické) a veřejné, státní a příspěvkové spíše méně



Důvody k nevytvoření DMP



„K čemu využíváte plán správy dat?“

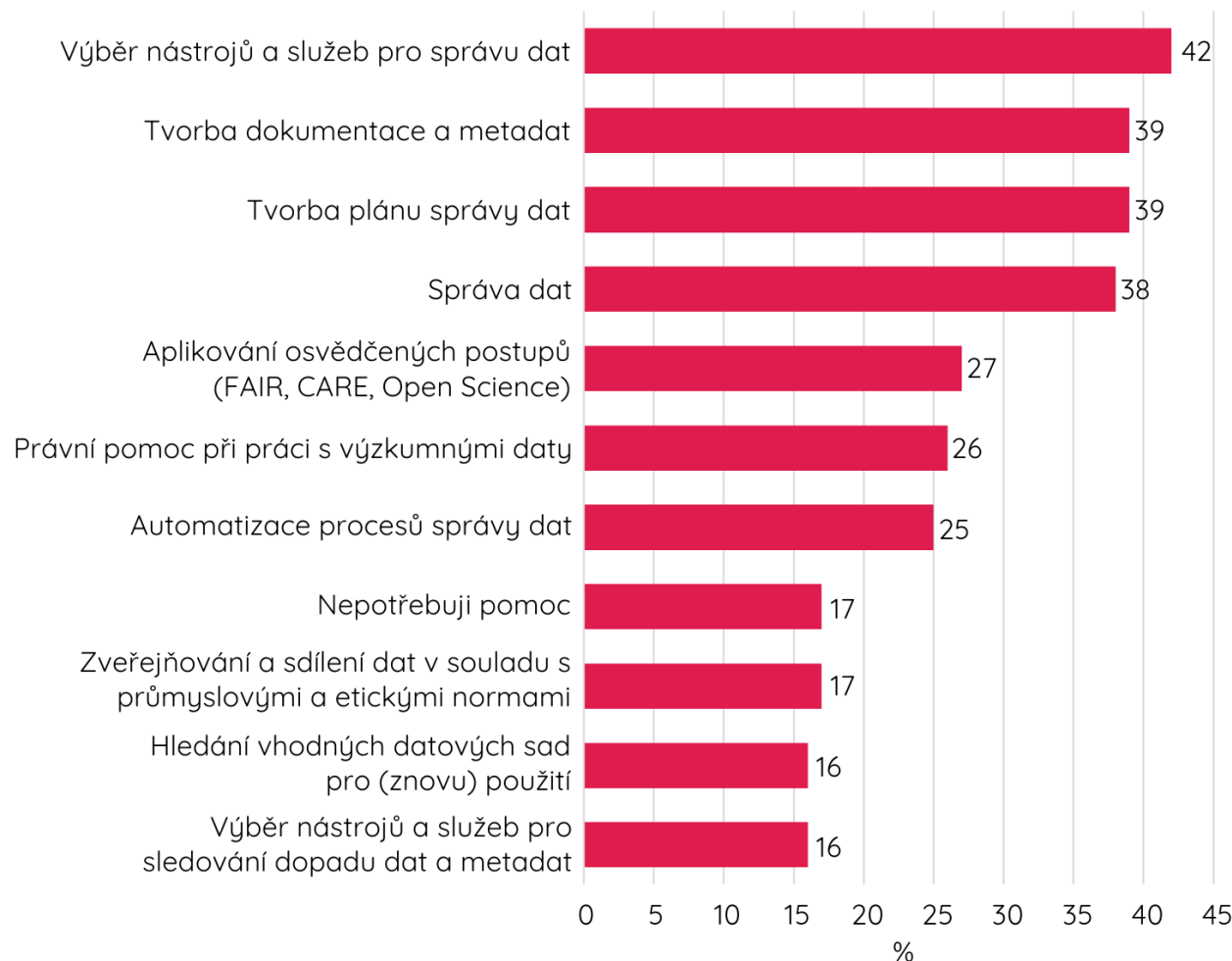


Spokojenost se správou dat

- 65 % respondentů spokojeno s aktuálním stavem správy dat na jejich projektu
- spokojenost snižuje bezradnost s daty
- není rozdíl ve spokojenosti podle subjektu spravujícího data (data steward, IT...)
- Důvody? Absence pomoci („*nechávejí to na nás*“), nedostatečné množství financí a personální zabezpečení („*Pro menší instituci je složité vychovat/platit si specialistu pro poměrně rozsáhlou agendu*“)



„V jakých oblastech správy výzkumných dat byste uvítal/a pomoc?“



Postoje vědců k otevřené vědě

- tři výroky o DMP
 - „plán správy dat je důležitou součástí výzkumné práce“
 - „plán správy dat je byrokratický požadavek“
 - „plán správy dat zvyšuje důvěryhodnost vědy“
- respondenti rozdělení na dva přibližně stejně velké tábory podporujících DMP a odpůrců DMP
- pozitivní nebo negativní vztah nesouvisí s typem instituce nebo s typem výzkumu
- pozitivní vztah – společenské a lékařské obory, negativní vztah – přírodní, zemědělské a veterinární obory
- pouze malá menšina má problém s FAIR daty



Shrnutí části Podpora a rozvoj kapacit

- Data stewardi málo dostupní
- Většina respondentů si spravuje svoje data sama nebo se obrací na kolegy
- 60 % vytvořilo nebo plánuje vytvořit DMP, hlavní motivace – povinnost
- Většina respondentů spokojena se správou dat
- Respondenti chtějí pomoci s výběrem nástrojů/služeb a vytvoření DMP
- Polovina respondentů odpůrci DMP, polovina podporující



Celkové shrnutí

- První celonárodní přehled správy dat ve vědeckých projektech v ČR.
- Rostoucí povědomí, ale velké rozdíly mezi obory a institucemi.
- Většina ukládá data neformálně (PC/externí média);
 - nízké využití institucionálních úložišť.
- Repozitáře využívány omezeně, výrazné rozdíly mezi disciplínami.
- Bariéry: neznalost, právní nejistoty, administrativa.
- Slabé využívání standardizovaných metadat.
- Velké výzvy u citlivých dat (hl. medicína).
- Potřeba: bezpečná úložiště, podpora, jasné postupy, oborová řešení.
- Výstupy = podklad pro EOSC CZ, NRP a národní strategie.
- Předpoklad opakování výzkumu na přelomu let 2027 a 2028 – možnost srovnání

Děkujeme za pozornost

martin.vavra@soc.cas.cz

tomas.cizek@soc.cas.cz

jan.klusacek@soc.cas.cz