

# Plán správy dat Ústavu přístrojové techniky Akademie věd České republiky

## 1. Shrnutí dat

V tomto plánu považujeme za „data“ relevantní k tomuto plánu pouze ty datové sady, které musí být zpřístupněny podle pravidel a principů otevřeného přístupu k vědeckým datům (dále „otevřená data“), data, která se povinně zpřístupňovanými daty podle aktuálních pravidel mohou stát v budoucnu, a data vytvořená v zájmu mimoústavních subjektů v rámci smluvního výzkumu nebo činnosti infrastruktur. Všechna ostatní data shromážděná nebo zpracovaná během výzkumné činnosti Ústavu přístrojové techniky AV ČR, v. v. i., (dále „ÚPT“) se považují za interní data, která jsou buďto nedůležité povahy (tj. nemusí být zpřístupněna pro podporu závěrů výzkumu), nebo se jedná o data, pro která existuje důvod zabraňující zveřejnění, tedy platná výjimka.

Vzhledem k technologickému a přístrojovému zaměření ústavu jsou významnou součástí výzkumu konstrukční práce v oboru elektrotechniky, strojního inženýrství, vakuové techniky aj. Výstupem, nezdědkou hlavním, jsou tedy data (informace) charakteru konstrukčních řešení. Může se jednat o strojírenské výkresy, elektrotechnická schémata, software atd. Vlastní data v obvyklém smyslu pojetí dat pak mohou být data naměřená prostřednictvím aparatur a jejich primární smysl spočívá v ověření funkčnosti a parametrů navržených přístrojových celků či subsystémů. Takováto data mají zásadní význam pro další přístrojový vývoj pracoviště ÚPT, jejich relevance pro externí zájemce je ovšem značně omezená, a to především v důsledku jejich původu ve specifických přístrojových platformách vlastní konstrukce. Jejich porovnatelnost a reprodukovatelnost na mimo pracoviště původu je tedy velmi limitovaná.

Formáty dat, s nimiž je pracováno na ÚPT, mohou být velmi různorodé, což reflektuje široký technologický záběr pracoviště. Data lze kategorizovat rámcově takto:

1. Data tradičních formátů zpracovatelná běžným, široce rozšířeným software. Do této kategorie spadají číselné a textové soubory, obrazová data v běžných formátech aj.
2. Data získaná proprietárním software, jenž je součástí speciálních, ale komerčních přístrojů provozovaných na ÚPT.
3. Data zaznamenaná přístroji konstruovanými na ÚPT a ukládaná ve formátu specifickém pro konkrétní aparaturu.
4. Data konstrukčního charakteru.

Využití dat v interním prostředí a jejich správa pro potřeby výzkumu řešitelských kolektivů jednotlivých projektů je ve správě oddělení, skupin a řešitelů.

## 2. FAIR data

### Zajištění dohledatelnosti dat (Findability), včetně ustanovení pro metadata

Metadata jsou nezbytnou součástí ukládaných datových souborů a nesou informace o době vzniku datového souboru, okolnostech jeho záznamu, typu experimentu a projektu, v jehož rámci řešení byla pořízena. Metadata související s jejich interpretací mohou v řadě případů zahrnovat prvky duševního vlastnictví a jejich zveřejnění či předání externím subjektům podléhá pravidlům pro zpřístupnění dat s aplikačním potenciálem (viz níže).

Vzhledem k širokému oborovému rozkročení ÚPT jsou další detaily struktury metadat specificky vázány na zvyklosti jednotlivých příslušných oborů. V tomto plánu správy dat nejsou a ani nemohou být jmenovitě upravována již jen pro variabilitu jednotlivých projektů a jejich specifických požadavků.

Data ÚPT nemají (až na výjimky) charakter dat sbíraných na souborech objektů či vzorků, které by měly relevanci pro externí strojové zpracování (data harvesting). Nebudou tedy takto pro tento účel upravována. Ve výjimečných případech na vyžádání mohou být vybrané datové soubory takto upraveny, pokud to bude externím subjektem vyžadováno pro smysluplný účel a nebude identifikováno riziko ohrožení duševního vlastnictví či kompromitace aplikačního potenciálu.

### Zpřístupnění dat (Accessibility)

Formy zpřístupnění dat ÚPT lze kategorizovat podle charakteru subjektu, kterému mají být k dispozici:

#### 1. Zpřístupnění pro řešitelský kolektiv ÚPT

Zpřístupnění dat pro interní potřeby řešitelských kolektivů projektů řešených na ÚPT je plné a neomezované. Přístup do datových úložišť je chráněn standardními obvyklými postupy zabezpečení dat. Metadata reflektují srozumitelnost interně potřebnou v rámci řešitelského kolektivu.

#### 2. Zpřístupnění pro pracovníky ÚPT mimo řešitelský kolektiv konkrétního projektu

Zpřístupnění dat pro osoby mimo konkrétní řešitelský kolektiv v rámci ÚPT je plné jako v bodě 1, pokud není řešitelský kolektiv smluvně vázán mlčenlivostí (NDA).

#### 3. Zpřístupnění pro členy konsorcia projektů, s nimiž ÚPT váže smlouva o spolupráci

V případě dat vzniklých v rámci projektů, které zahrnují externí spoluřešitele, jsou formy ukládání dat, charakteru metadat a formy sdílení dat určeny smlouvami s partnery, případně zadávacími dokumentacemi jednotlivých konsorciálních projektů. Tyto projektové a smluvní požadavky jsou v konkrétních projektech nadřazeny tomuto plánu správy dat.

#### 4. Zpřístupnění pro externí subjekty

Data, která jsou identifikována jako data, pro něž je zpřístupnění externím subjektům umožněno, jsou zpřístupněna v režimu "upon reasonable demand" – na základě smysluplného požadavku externího subjektu. Vlastní procedura předání požadavku na přenos dat a jeho uspokojení využívá webového formuláře na webových stránkách ÚPT.

Požadavek je směřován na osobu Data Stewarda, který/-á jej předá relevantnímu pracovníkovi ústavu. Rozhodnutí, jedná-li se o smysluplný ("reasonable") požadavek, náleží řešiteli projektu, v jehož rámci data vznikla. V případě potřeby řešitel konzultuje s vedoucím skupiny nebo oddělení.

Zpřístupnění dat externím subjektům není anonymní, požadavek na předání dat je vázán na vyplnění dotazníku zahrnujícího jméno, afiliaci a e-mailovou adresu žádajícího.

Formy zpřístupnění dat ÚPT lze kategorizovat také podle charakteru projektu, v jehož rámci vznikla, a jeho výstupů:

#### 1. Zpřístupnění dat souvisejících s publikacemi

Data související s publikacemi podléhají pravidlům vydavatelů a jsou zpřístupňována náležitou formou, vydavateli požadovanou.

#### 2. Zpřístupnění dat vzniklých v rámci projektů různých poskytovatelů

Taktéž zadávací dokumentace a pravidla poskytovatelů účelové podpory v podobě projektů stanovují vlastní pravidla práce s daty. Pokud tato pravidla nejsou v souladu s tímto plánem správy dat, pro konkrétní jednotlivé projekty jsou nadřazena tomuto plánu.

### **Repozitáře (úložiště)**

Důvěryhodnost datových úložišť a repozitářů je předmětem velké pozornosti a péče. Zásadním pro správu dat ÚPT jsou interní datová úložiště, která jsou součástí počítačové síťové infrastruktury ústavu. Péče o náležité zálohování, bezpečnost a spolehlivost provozu je povinností Oddělení informačních technologií ÚPT. Jednotlivá vědecká oddělení mají dedikován datový úložný prostor a využívají jej v režimu odpovídajícím vlastním potřebám. Další rozvoj a posilování interních datových úložišť jak co do zvyšování objemu uložitelných dat, tak co se týče posilování bezpečnosti, jsou jednou z priorit ÚPT.

Mimo interní datová úložiště využívá ÚPT řady dalších repozitářů a úložišť. Lze je kategorizovat takto:

1. Datové repozitáře oborové, v široké míře využívané mezinárodní komunitou s důvěryhodností danou šíří svého používání ve vědě. Kupříkladu arXiv, nebo ZENODO.

2. Důvěryhodné národní repozitáře nabízené národními infrastrukturami pro vědu a výzkum, jmenovitě EOSC-CZ, KNAV-ASEP, CESNET.

3. Repozitáře vydavatelů vědeckých publikací, jejichž použití pro data související s publikacemi je pro autory povinností.

4. Specifické oborové repozitáře dedikované pro užší vědeckou mezinárodní komunitu, jejichž využití je obvykle nutné v rámci řešení oborových konsorciálních projektů.

Využití jiných datových úložišť a repozitářů je možné po náležitém prověření jejich důvěryhodnosti a se souhlasem vedoucího příslušného vědeckého oddělení.

## Data

S ohledem na technologické zaměření ústavu má řada výsledků aplikační potenciál a/nebo je vytvářena přímo ve spolupráci s partnery z aplikační sféry. Existuje zde tedy možnost budoucí komercializace, případně převodu do praxe v širším smyslu. Přístup k datům je pak omezen nebo embargován tak, aby byla zajištěna ochrana duševního vlastnictví. Pochopitelně je nutné odepření přístupu k jakýmkoli výzkumným datům, která může konkurent použít k narušení ochrany IP nebo k zabránění zavedení ochrany (např. zabránění patentování nebo udělování licencí). Znamená to tedy, že všechna výzkumná data týkající se potenciálně komerčně využitelných a aplikovatelných výstupů jsou považována za data, která nespádají do kategorie výzkumných dat, na která se vztahují principy otevřeného přístupu.

Nezpřístupnění dat – data se nepřístupňují v těchto případech:

### 1. Data projektů aplikovaného výzkumu

Toto omezení se netýká dat souvisejících s publikacemi vytvořenými v rámci těchto projektů.

### 2. Data projektů smluvního a kolaborativního výzkumu financovaného z neveřejných prostředků a data vytvořená v rámci „Jiné činnosti“

### 3. Data s aplikačním potenciálem a data zahrnující prvky duševního vlastnictví

Rozhodnutí, zda datový soubor má aplikační potenciál či obsahuje prvek duševního vlastnictví, náleží řešiteli projektu, v jehož rámci data vznikla, konzultace s vedoucím skupiny či oddělení je hodna k doporučení.

### 4. Data vázaná smlouvami o mlčenlivosti (NDA)

### 5. Data konstrukčního charakteru – strojírenské výkresy, elektrotechnická schémata, software (s výjimkou software šířeného formou Open Source)

### 6. Data, která byla vytvořena proprietárním software vázaným na omezení licenčními ujednáními na tento software a související metadata

### 7. Data s etickými omezeními

### 8. Data, která narušují principy ochrany osobních údajů (GDPR)

## Interoperabilita dat (Interoperability)

Interoperabilita dat je pro jejich sdílení, výměnu a opakované používání zásadní. Primárním imperativem je zajištění interoperability v rámci řešitelských kolektivů pro interní použití. Proto jsou data opatřena adekvátními metadaty a srozumitelným popisem. Formáty a standardy odpovídají zvyklostem v oboru a potřebám řešitelských kolektivů. Řešitelům projektů je doporučeno používat standardní metadata, která navrhuje The Research Data Alliance, pokud není stanoveno v zadávací dokumentaci projektů jinak, nebo smluvně ve smlouvách o spolupráci na projektech. Pro všechny typy výzkumných dat generovaných v projektu budou používány standardní slovníky.

Interoperabilita dat relevantních k vědeckým publikacím bude z technického hlediska zajištěna uložením strojově čitelné elektronické verze zveřejněné publikace nebo konečné verze recenzovaného rukopisu přijatého ke zveřejnění do příslušných repozitářů. Z hlediska právního

zajištění otevřeného přístupu k výzkumným datům uloženým v těchto repositářích za podmínek poslední dostupné verze veřejné licence CC BY nebo jejího ekvivalentu.

### **Zajištění opakovaného použití dat (Reusability)**

Interní využitelnost dat je zajištěna uložením v interních datových úložištích, uložením ve formátech používaných standardně řešitelskými kolektivy, provozovanými laboratorními aparaturami, souvisejícím softwarem a opatřením adekvátními metadaty zajišťujícími srozumitelnost.

Data otevřená k přímému externímu využití jsou data vázaná na konkrétní výstupy publikačního charakteru v příslušných repositářích. Tyto zahrnují informace o dílčích výsledcích výzkumu nebo nástrojích a instrumentech potřebných k opětovnému využití výzkumných dat či k jejich validaci. Otevřený přístup k datům bude zajištěn v těchto repositářích po splnění podmínek poslední dostupné verze veřejné licence CC BY nebo jejího ekvivalentu.

Data určená ke sdílení v rámci konsorciálních projektů jsou dostupná členům konsorcia pro využití partnery projektů. Kvalita výzkumných dat bude zajištěna etickými kodexy vědeckého pracovníka platnými v institucích konsorcia.

## **3. Další datové výstupy výzkumu**

Jak již bylo popsáno v úvodu, technologické a přístrojové zaměření ústavu směřuje k výstupům, které mají charakter konstrukčních řešení, v podobě strojírenských výkresů, elektrotechnických schémat, software atd. Ve většině řešených projektů jsou tyto výstupy, včetně realizovaných aparatur, přístrojových celků, či subsystémů primárním typem výstupu, zatímco data sekundárním.

Správa hmotných i nehmotných výsledků konstrukčního charakteru a charakteru souvisejícího know-how jsou historicky předmětem náležité péče na ÚPT, hmotné výstupy jsou v laboratořích provozovány, udržovány a opravovány do doby, než dospějí do konce své životnosti.

Jejich sdílení externím subjektům je a může být upraveno řadou pravidel daných projekty, v jejichž rámci byly vytvořeny, smluvními ustanoveními ve vztahu k poskytovatelům a partnerům, jednalo-li se o výzkum ve spolupráci.

Zvláštní případ představují infrastruktury a jejich sdílení v režimu Open Access, který zajišťuje otevřený přístup k infrastrukturám a musí uživatelům zajistit i přiměřenou ochranu jimi poskytnutých i v ÚPT vytvářených dat v zájmu ochrany duševního vlastnictví uživatelů. Nakládání s těmito daty je upraveno specifickou smlouvou o předání dat.

## **4. Alokace zdrojů pro správu dat**

Zdroje alokované péči o data na ÚPT směřují především do investic do infrastruktury počítačové sítě a datových úložišť. Na pravidelný upgrade síťových, počítačových a datových prostředků je alokován adekvátní roční rozpočet. Nákupy investičního charakteru jsou podle potřeby zařazovány do plánu přístrojových investic ÚPT.

Lidské zdroje dedikované managementu dat: správa dat a péče o ně na úrovni projektů je úkolem řešitelů projektů. Interoperabilita dat na úrovni vědeckých skupin je svěřena vedoucím skupin a na úrovni oddělení vedoucím oddělení. Péče o počítačovou síť, datová úložiště ústavu, bezpečnost dat a zálohování dat je úkolem Oddělení informačních technologií. Role Data Stewarda, jehož

úkoly jsou definovány v kapitole “Zpřístupnění dat“, je svěřena pracovníkům/pracovnicím knihovny ústavu.

## 5. Bezpečnost dat

Bezpečnost dat ukládaných na interní datová úložiště je zajišťována péčí Oddělení informačních technologií ÚPT. Jedná se o zabezpečení před nepřátelskými útoky, před selháním techniky a před dalšími riziky ztráty dat.

Ukládání dat na externí datová úložiště a repozitáře je zabezpečeno volbou takových úložišť a repozitářů, které jsou důvěryhodné a náležitě certifikované.

## 6. Etika

Data vytvářená na pracovišti ÚPT jsou ve většině případů data o neživé přírodě a tato nepodléhají žádným etickým omezením.

Výzkumná data související s lékařským výzkumem (na ÚPT výzkum v oboru lékařské přístrojové techniky) podléhají režimu upravenému nařízením Evropského parlamentu a Rady (EU) 2016/679, zákonu č. 110/2019 Sb. o zpracování osobních údajů v platném znění a zákonu 101/2000 Sb. o ochraně osobních údajů v platném znění. Tato data proto nemohou být zveřejňována a sdílána. Mohou být však opětovně poskytnuta výzkumným a dalším subjektům v rámci EU, kteří se v minulosti podíleli na jejich sběru nebo zpracování. V souladu s platnou legislativou mohou být data dále poskytnuta právníkům či fyzickým osobám vykonávajícím např. kontrolní činnost.

Etickým omezením podléhají také data související s výzkumem na zvířecích modelech, který na ÚPT probíhá. Jednotlivé experimenty na zvířecích modelech podléhají schvalovacímu procesu Etické komise pro práci s laboratorními zvířaty. Nakládání s daty, která jsou výsledkem těchto experimentů, je stanoveno v příslušných plánech experimentů.