Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

cenia

NERP

Dataport

Uživatelská příručka

Česká informační agentura životního prostředí
1. října 2022

Obsah

Úvod 1
1. Popis aplikace 2
Hlavní funkce 2
2. Definice 3
Datová sada 3
Indikátor 3
Indikátorová sada 3
Štítky (tagy, klíčová slova) 3
Metadata 3
Definice DS/ Instance DS/ Datová sada/ Distribuce DS 3
Proces schvalovacího workflow pro datové sady 4
3. Role 5
3.1. Poskytovatel 5
3.1.1. Seznam datových sad 5
3.1.2. Odevzdání datové sady 6
3.1.3. Editace již odevzdané DS 10
3.1.4. Kontrola kurátora 10
4. Kontakty 11

Úvod

Česká informační agentura životního prostředí (CENIA) získává každoročně stovky datových sad od desítek poskytovatelů různými komunikačními kanály. Tato data a jejich metadata CENIA přijímala, spravovala, archivovala a zpracovávala manuálním způsobem bez možnosti systematické kontroly procesu. Některé části systému navíc nevyhovovaly požadovanému standardu otevřených dat DCAT-AP.

Aplikace Dataport (vstupní portál) je odpovědí na výzvy k zefektivnění, automatizaci a urychlení správy a zpracování dat. Aplikaci lze potenciálně rozšířit o další funkcionality i nové agendy.

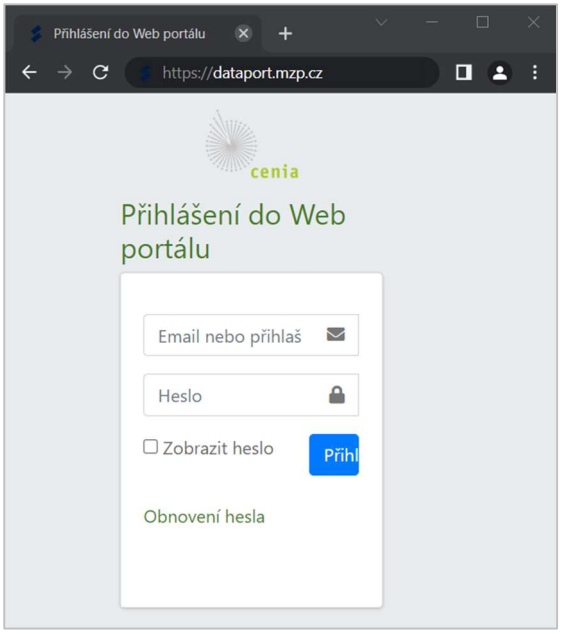
1. Popis aplikace

Dataport je aplikace nabízející uživatelsky přívětivé rozhraní pro ukládání a editaci datových sad, indikátorů, příslušných souborů, metadat a vazeb do datového skladu a řízení jejich publikace. Umožňuje větší kontrolu nad daty, dvoujazyčností a flexibilitu pro požadavky otevřených dat a publikaci dalších agend, jako např. SDG data.

Hlavní funkce

- Nabízí poskytovatelům dat nový komunikační kanál s přívětivým prostředím, validací a vzdálenou podporou kurátorů.
- Spouští skripty pro vytvoření katalogových listů v CKAN (instance datové sady a indikátoru) a vazeb mezi nimi.
- Nahrává datové sady a metadata do katalogových listů CKAN.
- Spouštění ETL a validačních skriptů.
- Správa uživatelů a bezpečnost – přístup pro externí poskytovatele datových sad, možnost měnit přístupové údaje.
- Integrace s CKAN a řízení CKAN přes RestAPI.
- Řízení publikace otevřených dat přes RestAPI (csv, json, metadata, doc, číselníky).
- Schvalovací workflow, notifikace.

Přístup do aplikace https://dataport.mzp.cz/ je povolen pouze uživatelům s přiděleným oprávněním, které administruje CENIA (role Správce).



2. Definice

Datová sada

Datová sada (DS) je vstupní soubor od poskytovatele dat opatřený metadaty. Odpovídá definici datové sady. Formát datové sady umožňuje transformaci do strojově zpracovatelného formátu. Slouží k napočítání indikátorů. Typicky se dále nepublikuje, ale na základě dohody mezi CENIA a poskytovatelem se může publikovat jak do portálu, tak jako Open Data. Instance datové sady odpovídá katalogovému záznamu ke konkrétnímu roku a obsahuje všechny zdrojové i transformované soubory a specifická metadata ke konkrétnímu roku.

Indikátor

Indikátor vzniká v CENIA a obsahuje data z 1..n datových sad. Všechna metadata zadává CENIA. Indikátor typicky obsahuje embed kód ke zveřejnění na webovém portálu a datovou sadu indikátoru ve formátu kompatibilním s formátem otevřených dat.

Indikátorová sada

Indikátorová sada je souhrn indikátorů patřících do stejné publikace.

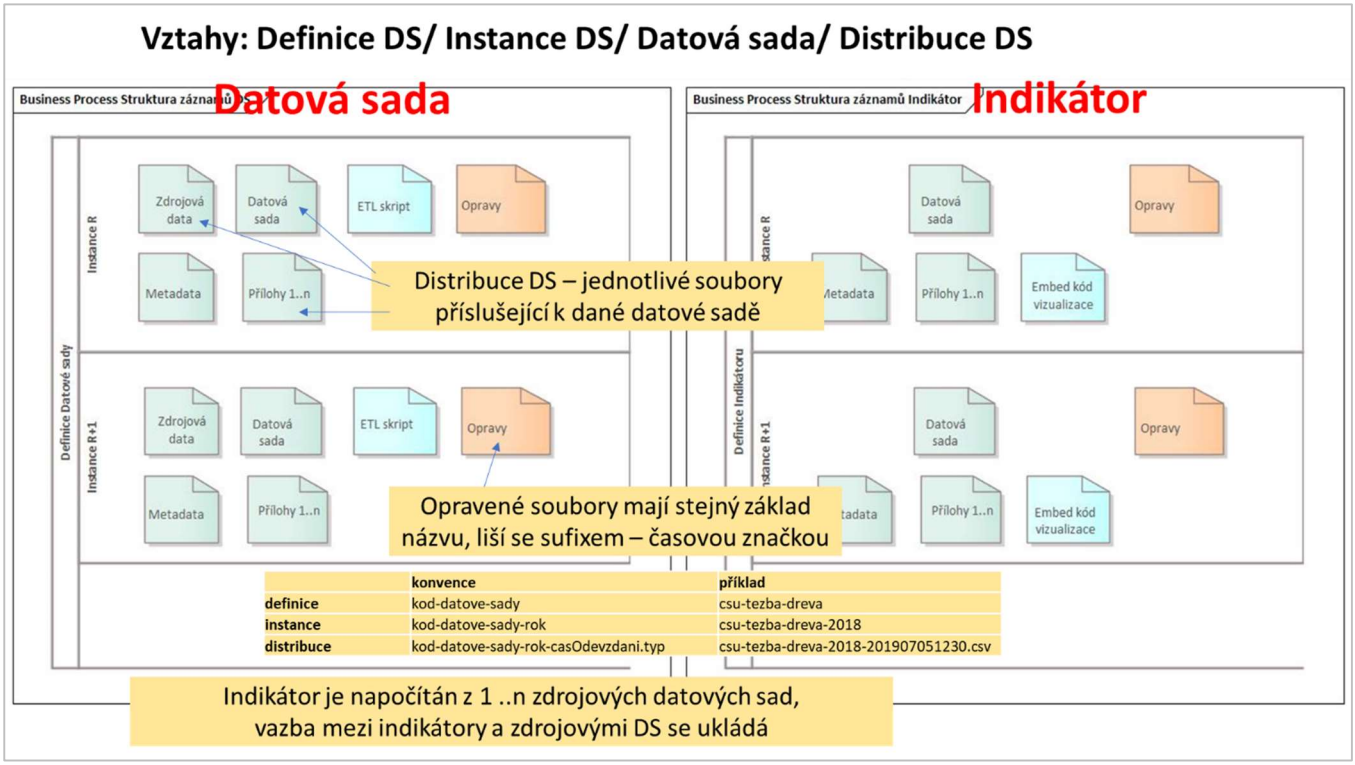
Štítky (tagy, klíčová slova)

Štítky se vybírají z klasifikačního slovníku GEMET. GEMET je specializovaný slovník pro oblast životního prostředí udržovaný organizací Eionet (EEA).
V případě publikace otevřených dat se štítky použijí pro položku „klíčová slova“.

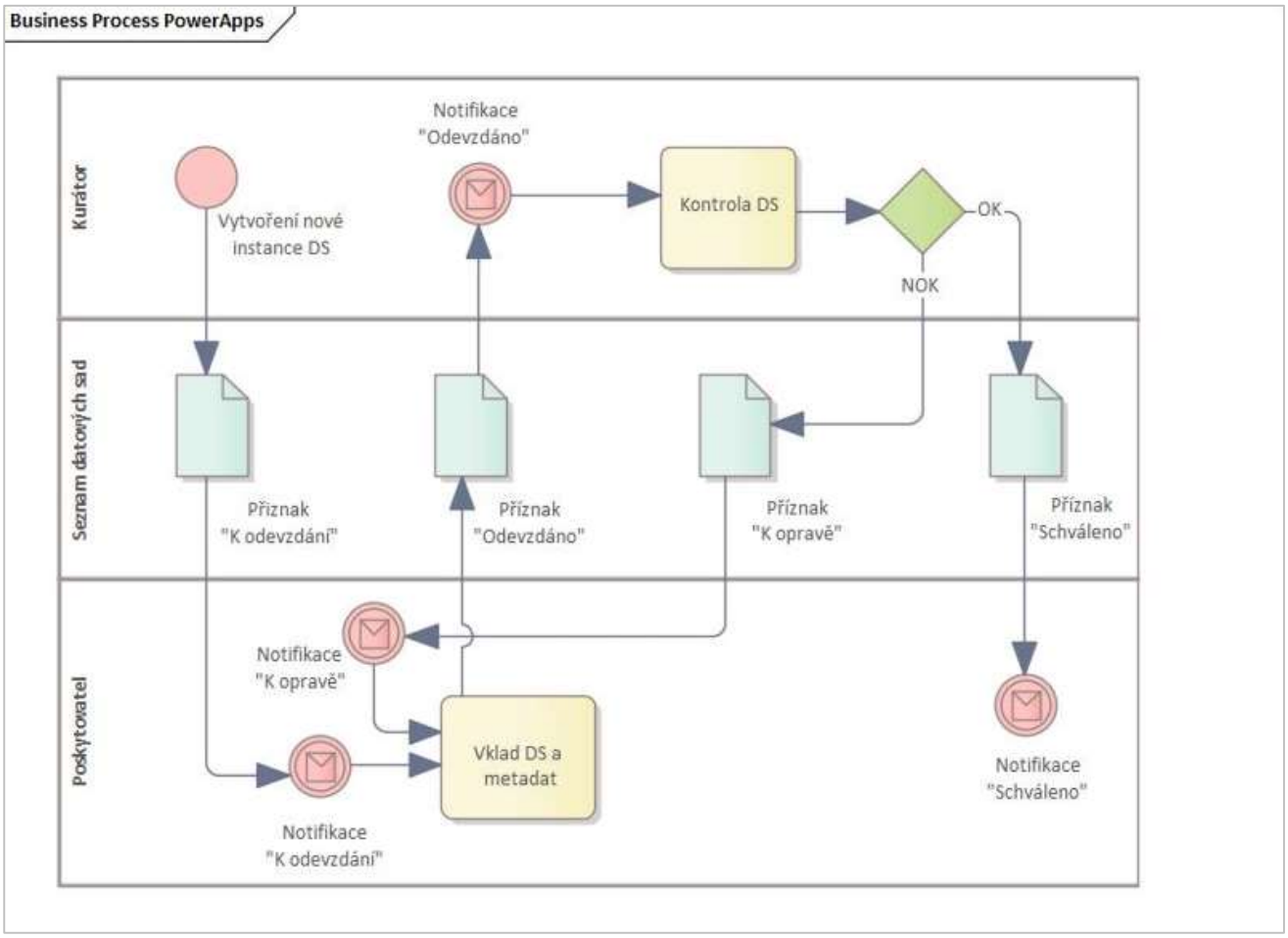
Metadata

Základní metadata vkládá Kurátor do definice DS. Metadata z definice DS Kurátor přenáší do instance DS pro jednotlivé roky.
Poskytovatel doplní jen chybějící metadata. Technická metadata (např. datumy) se generují strojově.

Definice DS/ Instance DS/ Datová sada/ Distribuce DS



Proces schvalovacího workflow pro datové sady



1. **kurátor** vytvoří **instanci datové sady** → **příznak „k odevzdání“**
2. **poskytovatel** odevzdá **novou datovou sadu** → **příznak “odevzdáno”**
3. **kurátor**
    a. **schválí** datovou sadu → **příznak “schváleno”**
    b. **neschválí** datovou sadu → **příznak “k opravě”**
4. **poskytovatel** odevzdá **opravu** → **příznak “odevzdáno”**
5. **kurátor**
    a. **schválí** datovou sadu → **příznak “schváleno”**
    b. **neschválí** datovou sadu → **příznak “k opravě”**

Kurátor může pro neaktivní datové sady nastavit → **příznak “neaktivní”**

3. Role

Systém rozlišuje tři uživatelské role: Poskytovatel, Kurátor a Správce.

	Poskytovatel	**Kurátor**	**Správce**
Oprávnění	podle příslušnosti k organizaci	libovolná datová sada	libovolná datová sada
Řízení přístupů			x
definice datové sady		x	x
definice indikátoru		x	x
založení instance datové sady		x	x
založení instance indikátoru		x	x
odevzdání datové sady	x	x	x
zadání a editace metadat datové sady	x	x	x
zadání a editace metadat indikátoru		x	x
výpis datových sad	x	x	x
schvalování datových sad		x	x
zneplatnění distribuce datové sady	x	x	x
změna kódu DS			x
dávková editace metadat			x
mazání záznamů/ definic			x

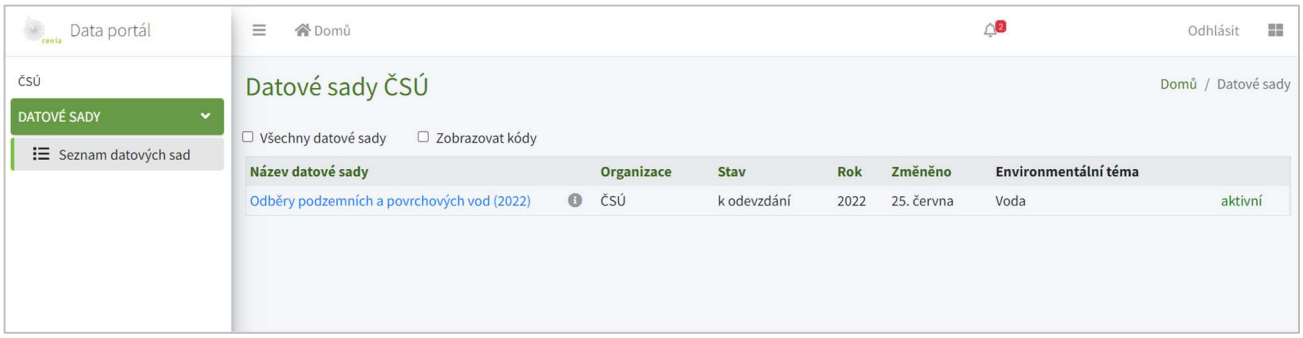
3.1. Poskytovatel

Základní oprávnění:

- Zobrazit seznam přiřazených datových sad (instance DS).
- Odevzdání dat.
- Oprava již odevzdané DS.
- Editace dat a metadat.
- Notifikace ve schvalovacím workflow.
- Aktivita všech uživatelů se loguje.

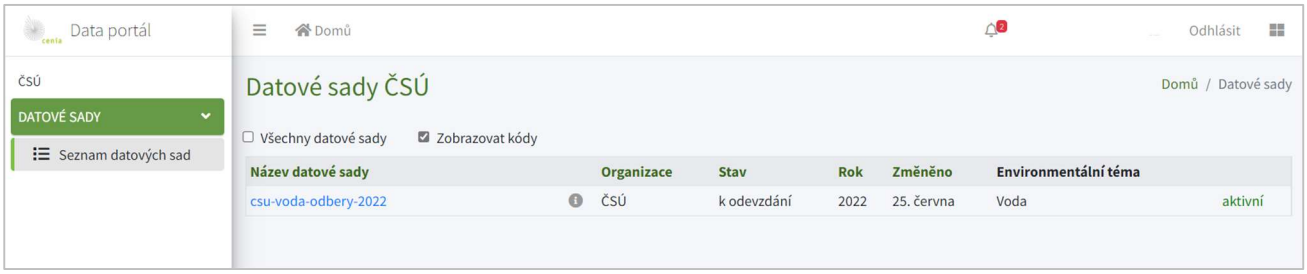
3.1.1. Seznam datových sad

Defaultně vyfiltrované jsou instance datových sad přiřazené jeho osobě.



Poskytovatel si může zaškrtnutím volby „Všechny datové sady“ zobrazit všechny instance datových sad jeho organizace. To mu umožňuje odevzdat DS i v zastoupení za kolegy ze stejné organizace. Každá změna je systémově logovaná.

S jednoznačnou identifikací DS může pomoci zaškrtnutí volby „Zobrazovat kódy“ místo názvu DS.



Zobrazené informace:

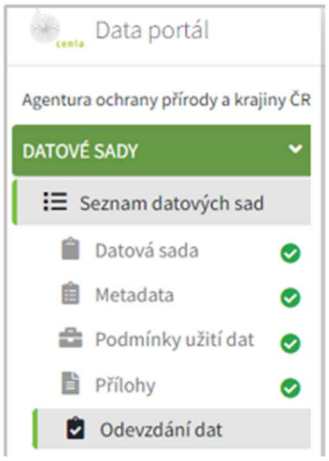
- Název datové sady (nebo kód instance DS pro daný rok).
- (náhled do definice DS).
- Organizace.
- Stav (k odevzdání, odevzdáno, schváleno, k opravě).
- Rok (instance DS).
- Změněno (datum poslední změny).
- Environmentální téma (Téma, kam je DS zařazená).
- aktivní/neaktivní (kurátor může konkrétní DS zneaktivnit).

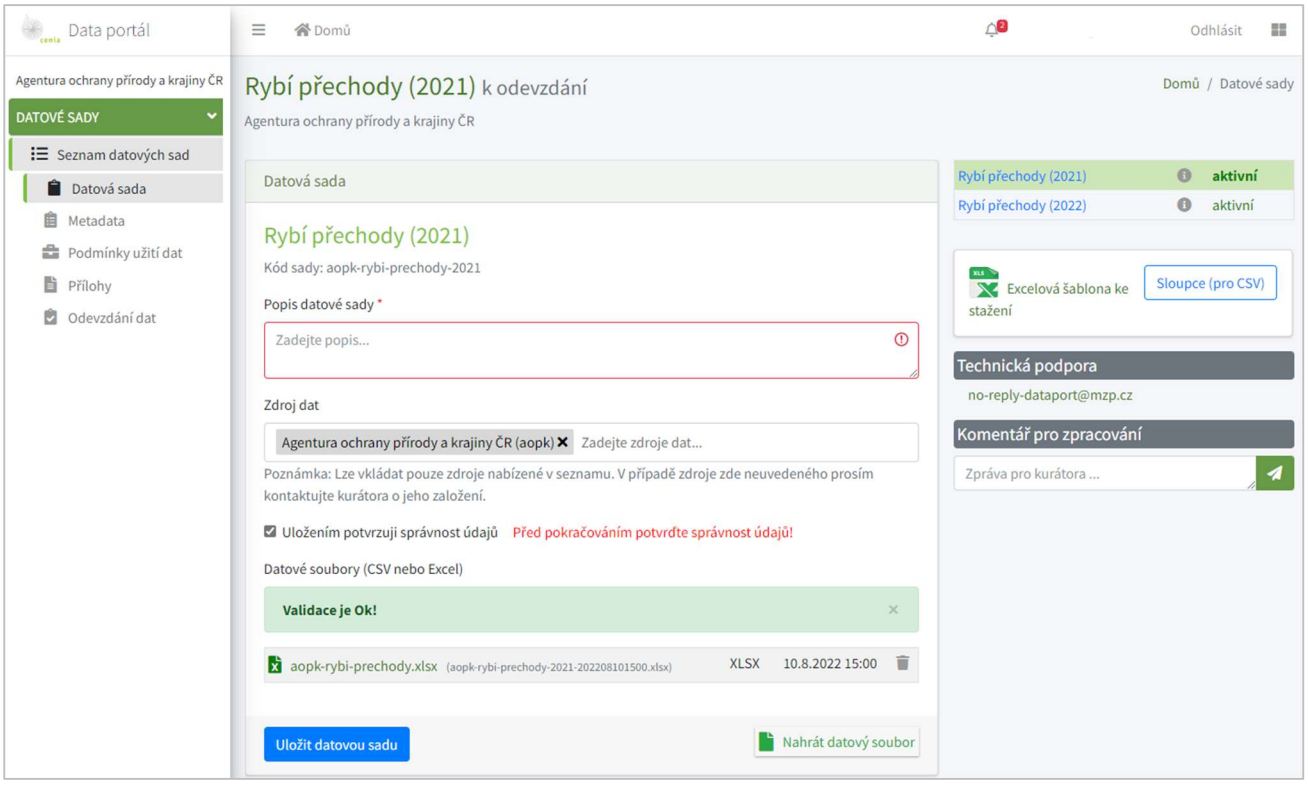
3.1.2. Odevzdání datové sady

Kliknutí na konkrétní instanci DS se otevře detail instance DS.

Záložka Datová sada

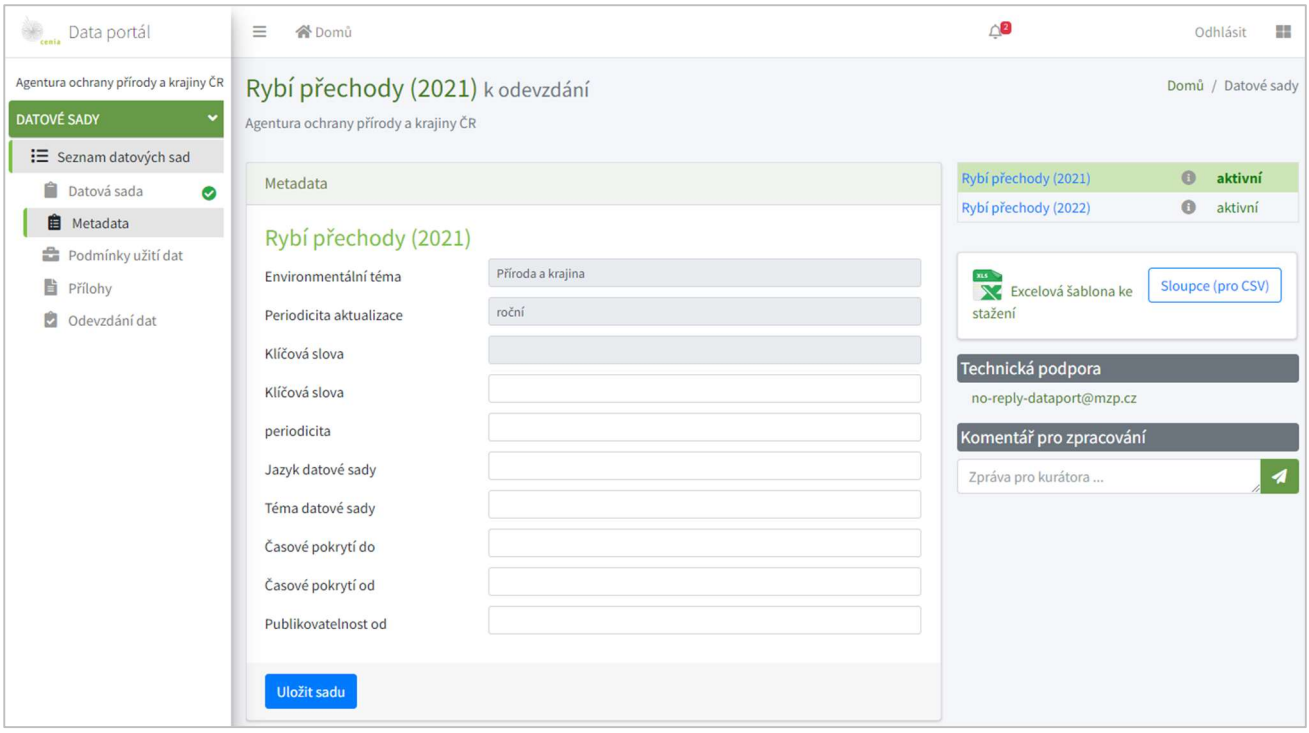
- Poskytovatel vyplní Popis datové sady.
- Je možné doplnit další zdroj dat.
- Tlačítkem „Nahrát datový soubor“ vloží datový soubor (formátu xls, xlsx, csv). V případě více listů excelu se spustí okno výběru konkrétního listu excelu, který chceme nahrát. Systémová validace soubor následně zkontroluje a oznámí, zda je soubor v pořádku. Je možné nahrát i soubor s chybou.
- Zaškrtne „Uložením potvrzuji správnost údajů“.
- Poskytovatel muže využít i „chatu“ s kurátorem v sekci „Komentář pro zpracování“, kam je možné vložit dotazy a poznámky ke zpracování. Odeslání zprávy v chatu je potřeba potvrdit tlačítkem .
- Vše uloží tlačítkem „Uložit datovou sadu“.
- Každá vyplněná záložka se po naplnění a uložení označí „bílou faječkou v zeleném kroužku“ .





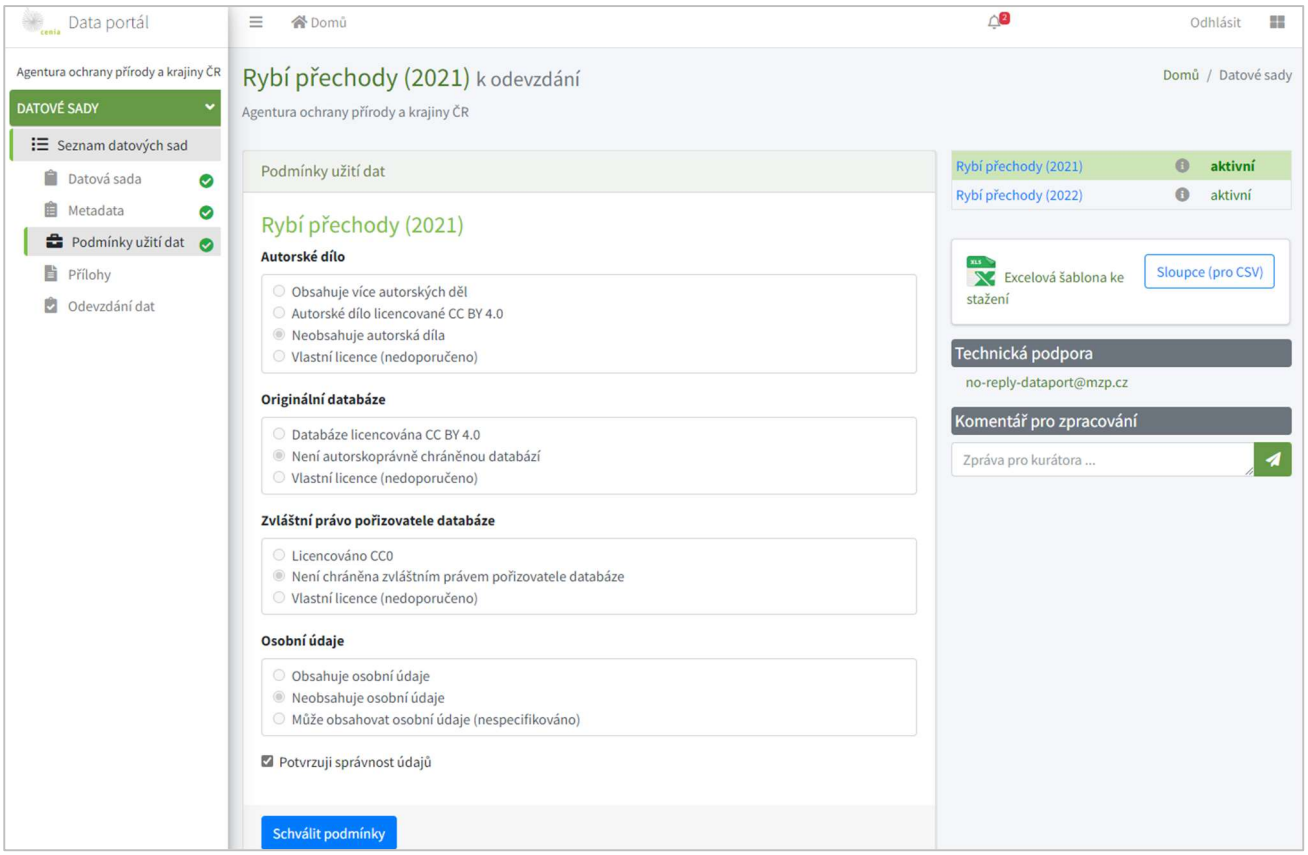
Záložka Metadata

- Poskytovatel vyplní průvodní informace k DS.
- Potvrdí tlačítkem „Uložit sadu“.



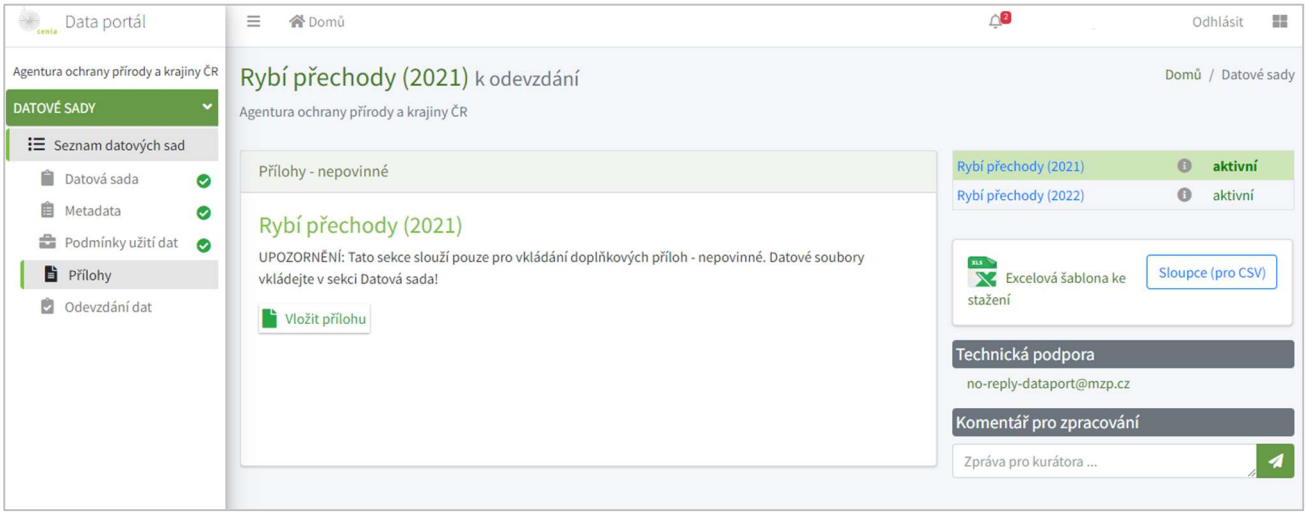
Záložka Podmínky užití dat

- Poskytovatel tlačítkem „Schválit podmínky“ odsouhlasí licenční podmínky.



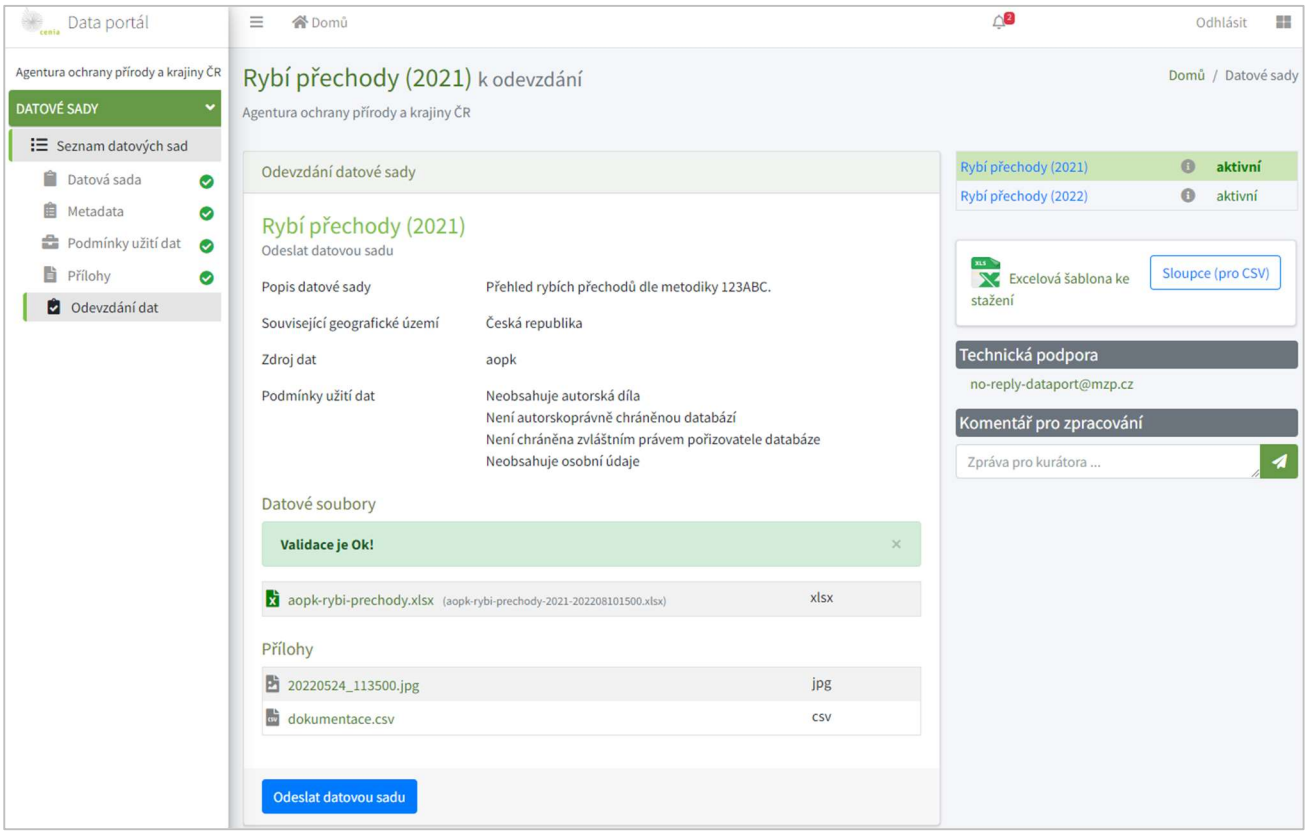
Záložka Přílohy

- Poskytovatel vloží průvodní dokumenty a přílohy (například metodiku, obrázky, mapy).



Záložka Odevzdání dat

- Poskytovatel zkontroluje sumarizaci informací o DS a tlačítkem „Odeslat datovou sadu“ celou DS pošle kurátorovi ke schválení.



Poskytovateli se zobrazí potvrzení o odeslání.

Nyní je možné vrátit se na Seznam datových sad, kde je již u odeslané DS stav „odevzdáno“.

3.1.3. Editace již odevzdané DS

Poskytovatel může editovat/opravovat pouze aktivní datové sady. Status aktivní/neaktivní určuje kurátor.

3.1.4. Kontrola kurátora

Odeslaná DS projde kontrolou od Kurátora, který ji udělí stav „schváleno“, nebo „k opravě“

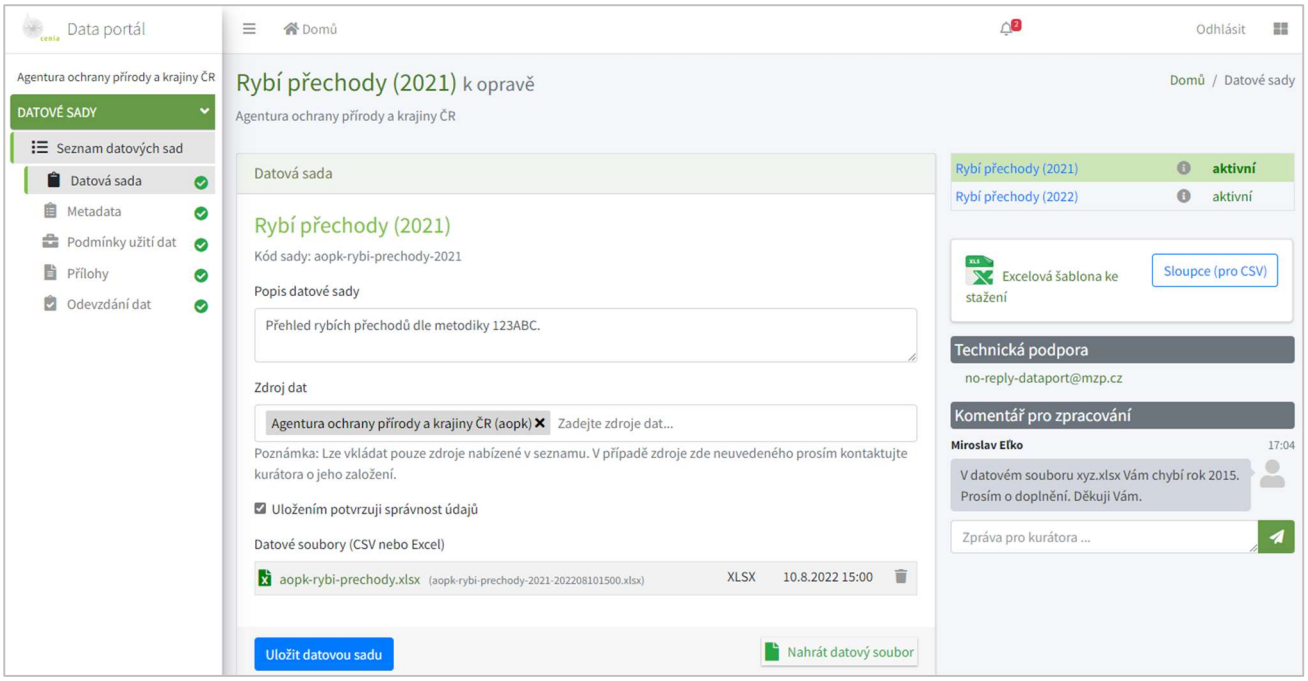
Schváleno

Stavem „schváleno“ kurátor potvrzuje poskytovateli úspěšné odevzdání DS.

K opravě

Datovou sadu ve stavu „k opravě“ je potřeba znovu otevřít a vyřešit připomínky/dotazy kurátora.

Pro sdílení informací k opravě využívá kurátor i poskytovatel chat „Komentář pro zpracování“.



Poskytovatel opravenou/doplněnou DS znovu odešle ze záložky „Odevzdání dat“ ke kontrole a schválení kurátorovi.

4. Kontakty.

Technickou podporu je možné kontaktovat emailem podpora-dataport@cenia.cz