



MANUÁL K SYSTÉMU THESES.CZ PRO VEDOUCÍ A Oponenty

Veřejně správní akademie – vyšší odborná škola, s. r. o.

Filipínského 1, 615 00 Brno

IČ: 25341901

IZO: 110030320

Platné od 1. února 2020

OBSAH

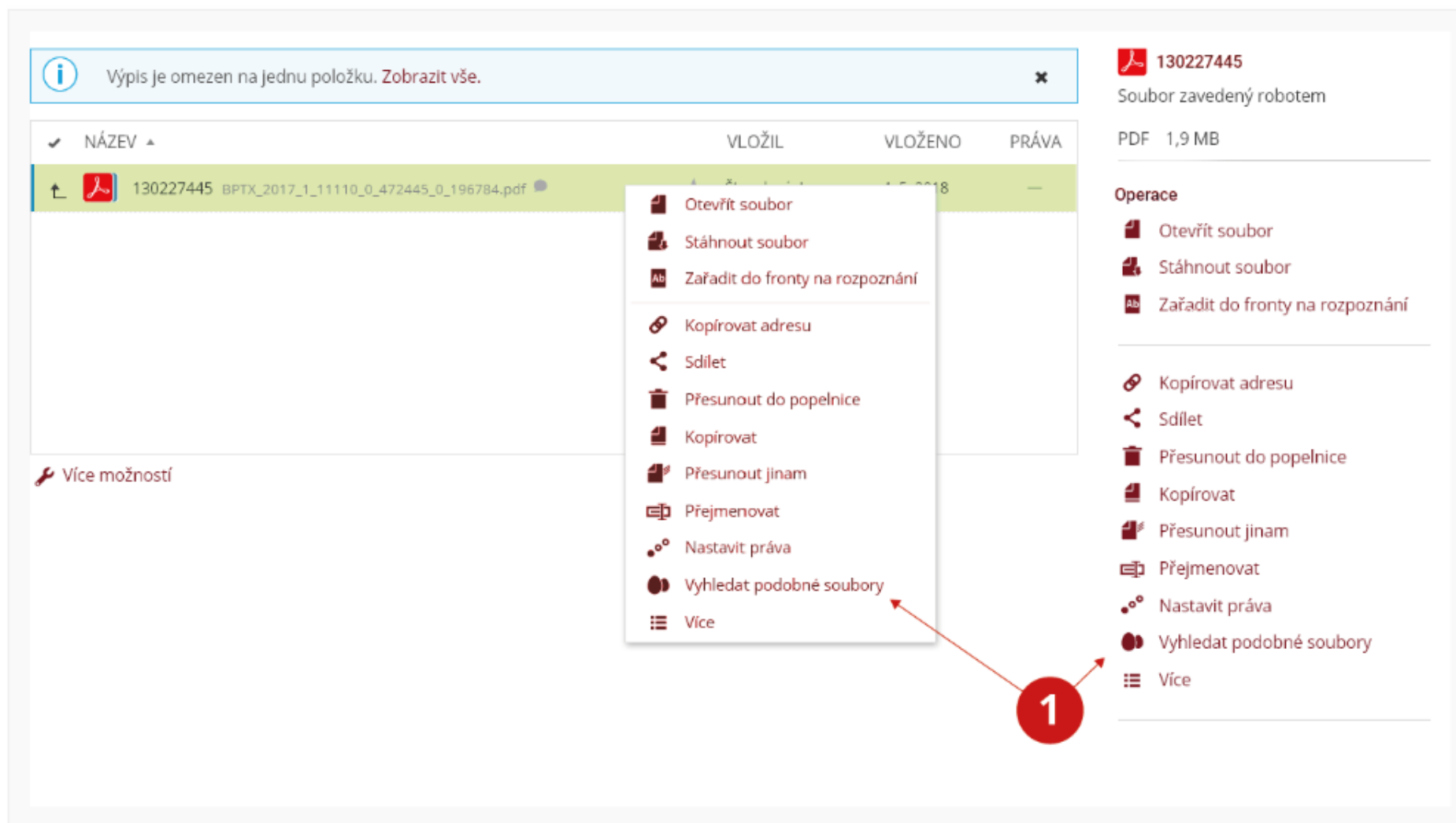
System theses.cz	3
Hledání podobností	4
Podobnosti s vybranými dokumenty	6
Podobnosti nejsou nalezeny	7
Označení souboru plagiátem	8
Porovnávání prací mezi sebou	8
Globální vyhledávání – možnosti podobností	9
Celková podobnost daného dokumentu	10
Podobnost mezi vybranými soubory	10
Podobnosti k jednotlivé práci	10
Výsledky stahování podobností	11
Zjistění stavu požadovaného dokumentu	11
Čtení souborů	12
System nastavení práv	13
SEZNAM OBRÁZKŮ	14

SYSTÉM THESES.CZ

- Do formuláře **student vyplní** povinná metadata.
- Po vyplnění formuláře se studentovi zobrazí archiv práce, kde vloží plný text své závěrečné práce (v podobě *.pdf, *.doc nebo *.txt).
- Plný text práce vloží **dostatečně brzy před obhajobou** (aby bylo možné zkontrolovat, zda se nejedná o plagiát).
- Archiv závěrečné práce je dostatečně naplněný, pokud:
 - ✓ anotace v češtině/slovenštině obsahuje minimálně 100 znaků,
 - ✓ anotace v angličtině obsahuje minimálně 100 znaků,
 - ✓ klíčová slova obsahují alespoň 30 znaků,
 - ✓ je zavedený plný text práce, který obsahuje alespoň cca 3000 znaků, a vytvořil se k němu textový soubor.
- Před obhajobami (odevdáním finální závěrečné práce k obhajobě) zkontroluje zpracovatel, zda je správně a dostatečně naplněn archiv závěrečné práce.
- Potvrdí její převzetí a od této chvíle se bude tato práce (metadata) zveřejňovat dle nastavení konfigurátoru a student už dál nemůže s archivem manipulovat (nemůže mazat, doplňovat text práce nebo měnit anotaci a ostatní metadata).
- Po obhajobě zadá zpracovatel datum obhajoby práce, anebo zadá „n“ nebo „-“ pokud práce nebyla obhájena. Pokud nebyla práce obhájena, zneprístupní se a záznam o práci bude moci vidět pouze uživatel s právem th_data_manager pro danou školu.
- K hromadným úpravám dat přijetí/obhajoby slouží aplikace „Theses.cz“ → „Správa dat“ → „manipulace se závěrečnými pracemi“, kde si zpracovatel může vyexportovat seznam ještě nepotvrzených prací (s naplněným archivem) nebo ještě neobhájených prací (ale již potvrzených a přijatých).
- Po kontrole, resp. doplnění data přijetí nebo obhajoby může tyto údaje hromadně importovat zpět do Theses.cz.

HLEDÁNÍ PODOBNOSTÍ

- Obsah souborů v datovém úložišti theses.cz, u kterých je k dispozici čistě textová verze, je průběžně strojově analyzován. Každou závěrečnou práci, kterou si v systému vyhledáte, **lze porovnat a zobrazit si soubory**, které mají podobný obsah:
 - ✓ kliknutím na název práce si zobrazíte základní informace o práci,
 - ✓ dole na stránce klikněte na řádek se souborem s plným textem práce (kliknutí pravým tlačítkem myši vyvolá panel podrobností vpravo, kliknutí levým tlačítkem vyvolá kontextové menu),
 - ✓ použitím operace „Vyhledat podobné soubory“ vyhledáte podobné soubory (viz *Obrázek 1*).
- **U nalezených podobných souborů systém vypíše míru shody v procentech a zobrazí červeně v textu jednotlivé podobnosti.** K prověření můžete vybrat jeden soubor či více souborů a porovnat si podobnosti jen s vybranými soubory. O globálním porovnávání souborů (tzn. všech vzájemně mezi sebou) se více dozvíte v nápovědě Globální vyhledávání plagiátů.
- Spočítání podobných dokumentů trvá jistou dobu. Zpravidla do jedné hodiny od zavedení nového dokumentu do systému je tento analyzován a připraven pro uživatelské hledání.
- Soubory se vyhledávají v celém systému, tedy bez ohledu na školu, kam práce patří.
- **Systém nezkontroluje seznam citací uvedený v dokumentu!**
- Porovnává se podle textových verzí souborů. V textové verzi nebylo nalezeno dostatečné množství slov, které by se pro porovnání mohlo použít. Nahlédněte do textové verze (s příponou `.txt`) souboru, zda se vytvořila správně. Pokud usoudíte, že se textová verze nevytvořila správně, můžete zkusit požádat vkladatele souboru, aby si sám ověřil (např. ve Wordu), zda lze soubor exportovat do textu; příp. aby zajistil nové vložení opravené verze. Problematické jsou HTML soubory, které obsahují hodně slov, které nejsou součástí textu (formátovací příkazy apod.). Také můžete poslat adresu chybného souboru na adresu správce theses@fi.muni.cz.
- **Údaj o podobnosti uvádí celkové procento podobnosti se všemi nalezenými podobnými dokumenty.**



1 Operace „Vyhledat podobné soubory“.

Obrázek 1 HLEDÁNÍ PODOBNOSTÍ

The screenshot displays a list of four document entries, each with a similarity percentage of '< 5 %' and a button to 'Zobrazit podobnosti pouze s tímto souborem'. The entries are:

- 1. Zdroj z Internetu, http://www.lf2.cuni.cz/Studium/disprohl.doc, Staženo: 25. 5. 2013. (Grey background, unchecked checkbox)
- 2. Zdroj z Internetu, http://www.adiktologie.cz/cz/articles/download..., Staženo: 19. 8. 2014. (Green background, checked checkbox)
- 3. Zdroj z Internetu, https://is.cuni.cz/webapps/zzp/download/13..., Staženo: 19. 4. 2015. (Grey background, unchecked checkbox)
- 4. Odevzdávárny, Barbora MÁCHOVÁ: UP Olomouc:2016/LS-KAZ/NSVKZ (2017). (Green background, checked checkbox)

At the bottom, a button labeled 'Zobrazit podobnosti s vybranými' is highlighted by a red circle with the number '2'.

- 1 Dokumenty, které chcete vyjmout z výpočtu podobnosti, odznačte.
- 2 Tlačítkem zobrazíte podobnosti pouze s vybranými soubory.

PODOBNOSTI S VYBRANÝMI DOKUMENTY

- Uživatel má možnost zaškrtnout dokumenty, pro které chce spočítat procento podobnosti zkoumaného dokumentu s vybranými podobnými dokumenty (viz *Obrázek 2*).
- Např. pokud se najde podobnost se zákonem, který je správně citován, uživatel může vyjmout tento dokument a přepočítat podobnost pouze s ostatními nalezenými podobnými dokumenty, ke kterým si zobrazí červeně zvýrazněné podobné pasáže.

Obrázek 2 PODOBNOST S VYBRANÝMI DOKUMENTY

Podobnost: **Psychologická výživa kojenců (2018)**

Celková podobnost: **1 %**

Jméno souboru: [Psychologická výživa kojenců \(2018\).pdf](#)

1

3 Celý text Jen podobnosti

2

4 Podobné soubory
Nápověda k podobnosti souborů
Podobnost s vybranými: **1%** podobnost s vybranými zdroji (4 z 4)

5 UPRAVIT VÝBĚR
 Nezahrnovat moje zdroje
 Bez zdrojů z internetu
 Jen více než % podobnost
 Jen v agendě Zdroj z Internetu
 Vybrat vše Odmítnout vše

6

7 Zobrazit podobnosti s vybranými

Podobné soubory

1 Zdroj z Internetu
<http://www.lf2.cuni.cz/Studium/disprohl.doc>
 Staženo: 25. 5. 2013
 < 5 %
 Zobrazit podobnosti pouze s tímto souborem

2 Zdroj z Internetu
<http://www.adiktologie.cz/c7/articles/download>
 Staženo: 19. 8. 2014
 < 5 %
 Zobrazit podobnosti pouze s tímto souborem

3 Zdroj z Internetu
<https://is.cuni.cz/webapps/zzp/download/13>
 Staženo: 19. 4. 2015
 < 5 %
 Zobrazit podobnosti pouze s tímto souborem

4 Odkazovaný
 : UP
 Olomouc:2016/LS-KAZ/NSVKZ (2017)
 < 5 %
 Zobrazit podobnosti pouze s tímto souborem

ABSTRAKT
 Tato bakalářská práce se zabývá... je rozdělena na část teoretickou a praktickou. V první části jsou rozebrána fakta týkající se novorozence, kojení a jejich potřeb. Popsány jsou základní fyziologické parametry novorozence a způsoby hodnocení stavu výživy. Práce sděluje aspekty kojení a umělé kojení a rozvedena je problematika zavádění nemléčných příkrmů. Také informuje o rizicích alternativního stravování kojenců, problému obezity a poruchách příjmu potravy. Část praktická je zaměřena na porovnání růstových grafů sta dětí v prvním roce života, z nichž je polovina kojených a polovina živěných umělými kojenými formulemi. Práce je zpracována kvantitativní analýzou dat. Z grafického zpracování vyplývá, že děti kojené mají nižší BMI a nižší průměrné váhové přírůstky než děti uměle živěné. Statistický test ale prokázal významné rozdíly jen u dívek v 6. měsíci.
 Klíčová slova: ...
 ABSTRACT
 It is divided into two parts. The theoretical part describes facts about newborn, infant and their needs. Then pursues basic physiological parameters of the newborn and the ways of assessment of nutrition status. The thesis discusses all aspects of breastfeeding, infant milk substitutes and complementary feeding. Also speaks about risks of alternative nutrition, obesity and eating disorders. The practical part finds out differences in growth charts of hundred children in the first year of their life. The half of the children are breastfed, the second half is artificially fed. This work is originated quantitative data analysis. The graphic processing shows that breastfed children have a lower body mass index and lower average weight gain. Statistical test demonstrates significant results only in 6th months old girls.
 Keys words: infant, nutrition, growth, breastfeeding, infant milk substitutes, complementary feeding, obesity
 Obsah
 1. Úvod.....9
 2. Demografická situace.....10
 3. Novorozenec.....11
 3.1. Vybrané charakteristiky fyziologického novorozence.....11
 3.2. Gastrointestinální trakt novorozence.....11
 3.3. Energetický metabolismus.....12
 4. Hodnocení stavu výživy.....12
 4.1. Růstové parametry.....12

- 1 Celková nalezená procentuální podobnost zkoumaného souboru vůči všem dokumentům v databázi.
- 2 Červeně zvýrazněné jsou podobné pasáže všech podobných dokumentů s daným souborem.
- 3 Lze si zobrazit jak celý text, tak třeba jen podobnosti.
- 4 Podobnost spočítaná pouze pro vybrané zdroje.
- 5 Omezení výběru zdrojů pro výpočet podobnosti (pro přepočítání je potřeba kliknout dole na tlačítko „Zobrazit podobnosti s vybranými“).
- 6 Zdroje, se kterými se u dokumentu nalezly podobnosti.
- 7 Jakmile vyberete pouze některé zdroje či jinak upravíte výběr zdrojů, je potřeba kliknout na na tlačítko „Zobrazit podobnosti s vybranými“.

Obrázek 3 PODOBNOSTI NEJSOU NALEZENY

PODOBNOSTI NEJSOU NALEZENY

- Podobnost, kterou systém našel mezi studentovou prací a prací/pracemi v databázi, nemusí vždy nutně znamenat, že se jedná o plagiát.
- Každou práci (podobnost) je nutné posoudit individuálně, zkontrolovat správnost citací apod.
- Neexistuje univerzální hodnota %, od které lze práci považovat za plagiát.

OZNAČENÍ SOUBORU PLAGIÁTEM

Následující informace na s. 8-11 jsou určeny zejména pro zjišťování podobností u problematických prací studentů.

- Označit je/není plagiát lze pouze při globálním vyhledávání plagiátů, které **mají k dispozici osoby s určitým právem na pracovišti**.
- Důvodem je, že není možné dopředu určit, zda dokument je/není plagiát bez toho, aniž by se prozkoumala podobnost se všemi ostatními soubory.
- Označení je/není plagiát se týká vždy dvojice dokumentů (nikoli pouze jednoho dokumentu). Toto označení je jednosměrné, tj. zkoumaný soubor je/není plagiátem k nalezenému podobnému souboru.

POROVNÁVÁNÍ PRACÍ MEZI SEBOU

- Jakmile se k nově vloženým závěrečným pracím vytvoří automaticky generovaná textová verze, lze je začít globálně porovnávat.
- Aplikaci globálního vyhledávání má k dispozici osoba s právem th_data_manager (tedy obvykle správce školy).
- V aplikaci: „Theses.cz“ → „Plagiáty“ → „Globální vyhledávání“ zvolíte Vaši školu, nejmenší procento podobnosti, které se má vyhledat (lze postupně snižovat dle počtu zobrazených souborů a vašeho posouzení, zda už systém nevyhledává zcela marginální podobnosti) a minimální počet znaků textové verze souboru, pro které se má vyhledávání provést.
- Dále je zde implicitně zaškrtnuta možnost „**ignorovat soubory závěrečných prací na adresách majících shodné dvě nejvyšší složky (/id/yyyyy/)**“. Soubory mající shodné nejvyšší složky typicky nebývají plagiáty.
- Vyhledávání můžete také omezit takto:
 - ✓ ignorovat dvojice souborů, které byly ručně označeny, že jsou plagiáty,
 - ✓ ignorovat dvojice souborů, které byly ručně označeny, že nejsou plagiáty,
 - ✓ ignorovat všechny soubory podobné souboru označenému „nikdy není plagiátem“,
 - ✓ omezit na soubory, které byly ručně označeny, že jsou plagiáty.
- Systém nyní vyhledává mezi všemi soubory (pracemi) v systému a zobrazí vám pouze ty soubory, které jsou podobné se soubory školy. Výsledkem vyhledávání je pak výpis zkoumaných souborů, u nichž systém vyhledal nějakou podobnost s jiným souborem/soubory. Procento podobnosti vyjadřuje, kolik procent textu zkoumaného souboru je podobných s textem nalezeného souboru/nalezených souborů. Výsledek vyhledávání obsahuje URL podobného souboru a na rozkliknutí i výpis textu, který je obsažen v obou podobných pracech. Pokud je soubor přístupný, zobrazí se také název práce a škola, které patří.

- Je-li u souboru uvedeno, které škole patří, kontaktujte správce školy pro Theses a domluvte se společně na dalším postupu.
- Je-li jako identifikace souboru uvedeno pouze URL, znamená to, že soubor není přístupný a po rozkliknutí URL zjistíte přímo kontakt na správce školy, kterému můžete URL zaslat (jemu se soubor zobrazí a zašle Vám ho).
- **Zkontrolujte citace a posuďte, zda se skutečně jedná o plagiát.**
- Výsledek vlastního porovnání můžete i s komentářem, zda se jedná/nejedná o plagiát, uložit např. do HTML souboru v dokumentovém serveru do složky školy.

GLOBALNÍ VYHLEDÁVÁNÍ – MOŽNOSTI PODOBNOSTÍ

- Máte dvě možnosti, jak získat protokol o nalezených podobnostech ve všech pracech školy – *manuální a automatická*.
 - ✓ *Manuální:* Po vyhledání podobností v aplikaci globálního vyhledávání „Theses.cz“ → „Plagiáty“ → „Globální vyhledávání“ si tyto informace můžete stáhnout v PDF protokolu nebo XML záznamu.
 - ✓ *Automatické:* Informace o globálním vyhledávání můžete stahovat ze systému také automaticky. Nejprve je nutné si vytvořit neosobní účet, pomocí kterého budete přistupovat k aplikaci. Neosobní účet může vytvořit pouze osoba s právem „th_data_manager“. Tato osoba bude za tento účet nadále zodpovědná. Informace můžete poté stáhnout pomocí skriptu v aplikaci „Theses.cz“ → „Plagiáty“ → „Globální vyhledávání“. Skript má jako povinný parametr „fak“, takže je třeba upřesnit, na které škole/fakultě chcete globální vyhledávání provést.

Příklad použití:

```
curl -u 200425:ti54as2341a "https://theses.cz/auth/dok/plag_th_glob_vyhledat.pl"
```

Další možné parametry:

from - jen práce do uvedeného data (ve formátu 'DD MM YYYY' nebo 'YYYY-MM-DD')

proc - jen podobnosti, které mají minimálně danou procentuální shodu (0-100)

vel - jen soubory, které mají minimálně daný počet znaků (např. 40 000).

ist2 - ignorovat soubory závěrečných prací na adresách majících shodné dvě nejvyšší složky (/id/yyyyyy/)

ijpl - ignorovat dvojice souborů, které byly ručně označeny, že jsou plagiáty

inpl - ignorovat dvojice souborů, které byly ručně označeny, že nejsou plagiáty

ukpl - omezit na soubory, které byly ručně označeny, že jsou plagiáty

pdf - místo XML záznamu dostanete PDF protokol

Obrázek 4 PŘÍKLAD POUŽITÍ

CELKOVÁ PODOBNOST DANÉHO DOKUMENTU

- Po zjištění podobnosti u vybrané práce je v xml dokumentu uveden i prvek `plg:doc_hash`, který jednoznačně identifikuje soubor v Theses.
- Pro zjištění celkové podobnosti lze použít odkaz: `https://theses.cz/auth/podob/XXX/?all+xml`
- Celkové procento je v prvku `plg:fullscore`, všechny podobné pasáže pak v prvku `plg:common`.
- Pro výstup ve formátu PDF stačí v odkazu nahradit atribut xml atributem pdf.

PODOBNOST MEZI VYBRANÝMI SOUBORY

- Pro zobrazení podobností ve vzorovém dokumentu 00000 s dokumenty 11111 a 22222 lze použít odkaz:
 - ✓ `https://theses.cz/auth/podob/00000/11111/22222/`, kde první identifikátor je vzorový soubor a další dva jsou identifikátory souborů, ke kterým chceme zobrazit podobnosti.
- Při zobrazení většího počtu podobných dokumentů, může být URL příliš dlouhé. V takovém případě lze podobnosti získat odesláním GET požadavku, který má následující parametry:
 - ✓ vzor: `doc_hash` vzorového dokumentu
 - ✓ dok: `doc_hash` podobného dokumentu

PODOBNOSTI K JEDNOTLIVÉ PRÁCI

- Podobnosti si můžete nechat zobrazit opět *manuálně* nebo *automaticky*.
 - ✓ *Manuálně*: v aplikaci „Theses.cz“ → „Plagiáty“ → „podobnosti k VŠKP“ si vyberte školu, ke které příslušná práce náleží. Dále vepište lokální identifikátor závěrečné práce (tzn. identifikátor, pod kterým je práce uložena v počítačovém systému poskytovatele).
- Na základě identifikátorů `pts:sender.id` (vybraná škola) a `pts:thesis.id` (resp. `dc:identifier` pokud používáte formát EVSKP) dostanete k dané práci výpis podobných prací a podobností buď v XML nebo PDF protokolu.
 - ✓ *Automaticky*: podobnosti k dané práci je možné volat vzdáleně podobně jako u globálního vyhledávání přes neosobní účet pomocí skriptu na stránce https://theses.cz/auth/plagiaty/plag_vskp.pl.

VÝSLEDKY STAHOVÁNÍ PODOBNOSTÍ

- Celkové procento podobností lze získat nebo vynechat z nalezených podobností dokumenty, které mají stejného autora jako vzorový dokument.
 - ✓ itos – ignorovat dvojice souborů, které mají stejného autora
 - ✓ all_perc – zobrazit celkové procento podobností u daného dokumentu
 - ✓ perc_min – zobrazit pouze podobnosti, které jsou větší než určené minimální procento (např.: perc_min=5 zobrazí podobnosti jen se soubory, jejichž podobnost je větší než 5 %)
 - ✓ pdf – výstupní formát ve formátu PDF
 - ✓ xml – výstupní formát ve formátu XML

ZJISTĚNÍ STAVU POŽADOVANÉHO DOKUMENTU

- Stav požadovaného dokumentu můžete zjistit prostřednictvím atributu plg:status, který se nachází v prvku plg:info.
- Atribut plg:status může nabývat hodnoty od 1 do 9 následovně:
 1. Dokument není v metadatech – neznámé url. Je nutné provést nový import.
 2. Dokument je připraven ke stažení (stahování by mělo proběhnout v noci).
 3. Dokument je zaveden v systému, ale nemá vytvořenou textovou verzi (může se zobrazit i důvod, proč bylo vytvoření neúspěšné nebo upozornění, že soubor není v textovém formátu ale např. ve formátu .eps apod.).
 4. Dokument není zkontrolován systémem na odhalování plagiátů – po vytvoření textové verze chvíli trvá, než dojde k nalezení podobností (prvek plg:info se zde nachází v prvku plg:record).
 5. K souboru nebyly nalezeny žádné podobnosti (prvek plg:info se zde nachází v prvku plg:record).
 6. K souboru byly nalezeny podobnosti, informace najdete ve vnořených prvcích plg:plagiat.
 7. Dokument zkontrolován u předchozí verze, podobnosti se teď přepočítávají.
 8. Dokument nezkontrolován, je příliš malý.
 9. Dokument nezkontrolován, obsahuje různá kódování textu.

ČTENÍ SOUBORŮ

Právo čtení souboru lze nastavit pro kohokoliv v internetu, pracovníky školy, aktuální studenty fakulty, osoby (zadáním konkrétního učo), výčet učo osob (zadáním více učo naráz) a absolventy (školy, fakulty). Měnit tato práva může obvykle správce školy. Informaci o tom, jaké právo soubor má, zjistíte podržením myši na ikonce ve sloupci „Práva“.

✓	NÁZEV ▲	VLOŽIL	VLOŽENO	PRÁVA
	Archiv závěrečné práce ESF N-FU FIN if4be /6		16. 9. 2014	
<input type="radio"/>	Posudek vedoucího posudek_vedouciho_... docx		24. 1. 2017	
<input type="radio"/>	Posudek oponenta posudek_oponenta_... docx		22. 1. 2017	
<input type="radio"/>	Plný text práce Financni_analyza_vybraneho...			Čte kdokoliv v Internetu
<input type="radio"/>	Klíčová slova keywords.txt		4. 1. 2017	
<input type="radio"/>	Anotace česky annotation.txt		4. 1. 2017	
<input type="radio"/>	Anotace anglicky annotation_english.txt		4. 1. 2017	

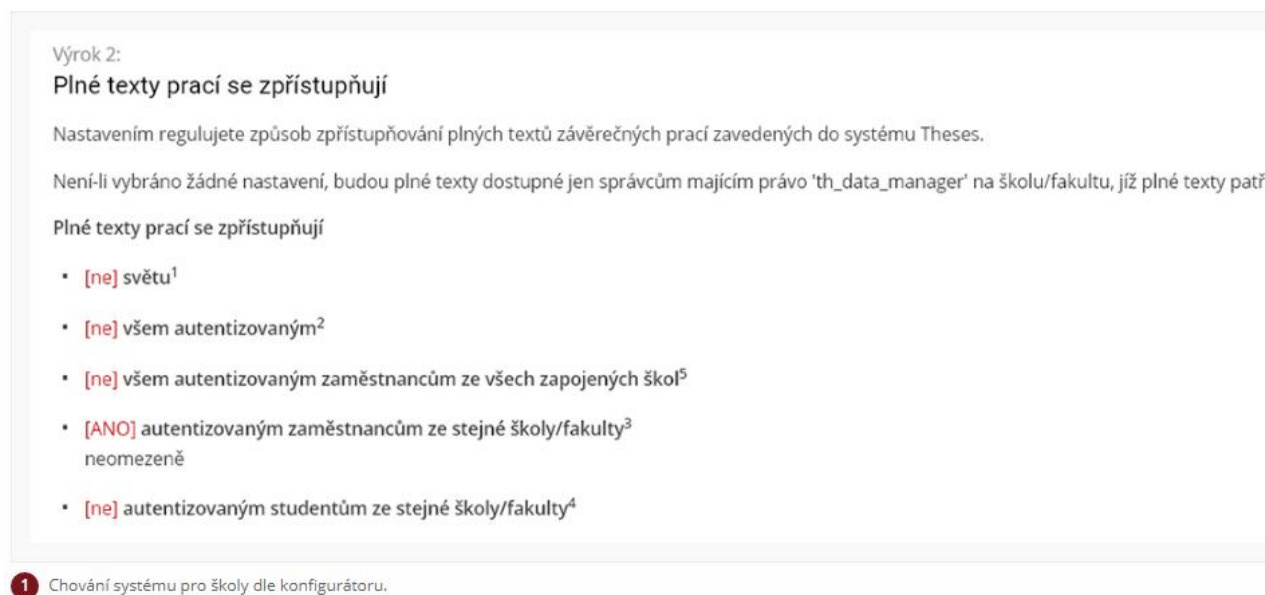
Více možností

1 Informace o právech souboru.

Obrázek 5 PRÁVA PRO ČTENÍ SOUBORŮ

SYSTÉM NASTAVENÍ PRÁV

- V systému Theses.cz fungují kromě práv pro konkrétní soubor ještě práva nadřazená, která platí pro větší celky (např. školu, fakultu) a řídí se tzv. Konfiguratorem: „Theses.cz“ → „Konfigurátor“
- Pokud je konfigurátor nastaven restriktivněji než práva u souboru, nelze souboru nastavit právo pro čtení větší, než povoluje Konfigurátor (viz *Obrázek 6*).



Obrázek 6 SYSTÉM NASTAVENÍ PRÁV

- I práva povolená „Konfiguratorem“ ale nemusí platit, je-li „Konfigurátor“ zneplatněn (žádná metadata ani závěrečné práce nelze číst). Systém tedy při otevírání souboru postupuje takto:
 - ✓ Zkontroluje, zda „Konfigurátor“ není zneplatněn, pokud ano, nelze soubor číst.
 - ✓ Zkontroluje, zda „Konfigurátor“ povoluje soubor číst.
 - ✓ Zkontroluje, zda práva u souboru jsou nastavena tak, že lze soubor číst.

SEZNAM OBRÁZKŮ

Obrázek 1 HLEDÁNÍ PODOBNOSTÍ	5
Obrázek 2 PODOBNOST S VYBRANÝMI DOKUMENTY	6
Obrázek 3 PODOBNOSTI NEJSOU NALEZENY	7
Obrázek 4 PŘÍKLAD POUŽITÍ	9
Obrázek 5 PRÁVA PRO ČTENÍ SOUBORŮ	12
Obrázek 6 SYSTÉM NASTAVENÍ PRÁV	13