

Řízené slovníky a ontologie v ČR

Jak fungují, kde je najít

Josef schwarz, cit ff mu, [lindat/clariah-cz](https://lindat.cz/clariah-cz)

26. 2. 2024



Spolufinancováno
Evropskou unií



MUNI
ICS

Registrační číslo IPs EOSC-CZ
CZ.02.01.01/00/22_004/0007682

- ARM (1)
- armáda (9)
- armáda (7)
- Armáda České republiky (4)
- Armáda České republiky (3)
- Armáda ČR (3)
- armáda hliněná (1)
- armáda terakotová (1)
- armádní (4)
- armádní generálové (8)
- armádní jednotky (3)
- armády (440)
- Armand, Inessa (1)

[Top of the AAT hierarchies](#)

.... [Objects Facet](#)

..... [Visual and Verbal Communication \(hierarchy nam](#)

..... [Visual Works \(hierarchy name\)](#)

..... [visual works \(works\)](#)

..... [<visual works by material or technique>](#)

..... [ceramic ware \(visual works\)](#)

..... [pottery \(visual works\)](#)

..... [stoneware \(pottery\)](#)

..... [black basaltes](#)

10.52 Zmrzlina, vč. subdodavatelských prací

1052100002 Zmrzlina a podobné výrobky (včetně ovocné dřeně a nanuků, kromě směsí a základů pro výrobu zmrzliny)

1052990011 Práce výrobní povahy ve třídě 1052

10.61 Mlýnské výrobky, vč. subdodavatelských prací

1061110002 Loupaná rýže (rýže cargo nebo hnědá rýže)

1061123002 Poloomletá nebo celoomletá rýže, též leštěná nebo hlazená

1061125002 Zlomková rýže

1061210002 Pšeničná mouka nebo mouka ze souřeže

1061220002 Obilné mouky, kromě pšeničné mouky nebo mouky ze souřeže

1061230002 Mouka, krupice a prášek ze sušených luštěnin, sága, manioku nebo ořechů a ovoce

1061240002 Směsi a těsta pro přípravu pečiva, cukrářských výrobků

ATQ	ダンス 一緒に使用：(すべてのATQ *コード) STYLE 修飾子 6 *必要に応じて
ATQC	振り付け
ATQL	バレエ
ATQR	社交ダンス 【適用可能】ラテンダンス
ATQT	コンテンポラリー・ダンス

potravinářství

Anglický ekvivalent

☉ food industry

Podřazená hesla

➤ bezpečnost potravin

➤ potravinářská biochemie

➤ potravinářská mikrobiologie

➤ potravinářské technologie

➤ potravinářský průmysl

➤ potraviny

Příbuzné koncepty

☉ Wikipedia: [Potravinářství](#)

learning difficulty

MT 02 learning
NT1 acalculia
NT1 dysgraphia
NT1 dysorthography
NT1 reading difficulty
NT2 dyslexia

learning habit

MT 02 learning
MT 23 behaviour incentive
BT1 habit
BT1 learning process
BT2 learning

learning kit

MT 08 equipment
BT1 multipurpose equipment
BT2 equipment

learning module

MT 02 learning
MT 09 curriculum
BT1 curriculum
BT1 learning strategy
BT2 learning process
BT3 learning
RT modular training

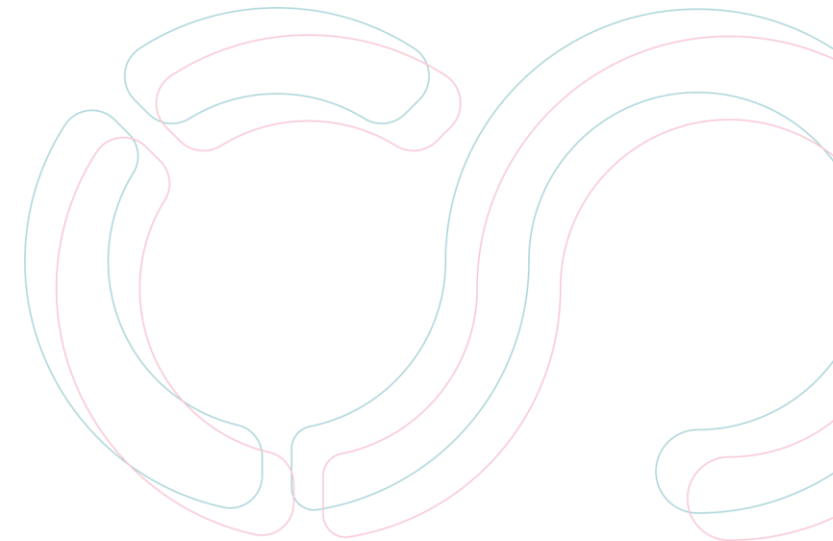
learning pace

MT 02 learning
BT1 time-factor

PN Christianity
 * <i>The normal retroactive principle for building classmarks is interrupted in PN to provide for special expansions, especially of Christology PNB and the practice of the Christian religion PNF/PNG. The overall order, however, remains consistent with that of PB/PD. [Suggested addition. Tony (1999)]</i>
PNA Theology (general)
 <u>Add to PNA letters A/Y following PA in PAA/PAY</u>
PNA Z Systematic theology
 <u>Add to PNA Z letters A/H following PB in PBA/PBH</u> e.g., Anthropomorphism PNA ZH
PNB Incarnation : Christology
PNB B Person of Christ
BH Historicity, Historical Jesus
BJ Names of Christ
BK Nature of Christ
BL Humanity
BM Divinity
CB Activities of Christ
CD Offices of Christ

Osnova

- Teoretická část
- Praktické ukázky
- Dotazy, debata



Teoretická část



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

MUNI
ICS

Registrační číslo IPs EOOSC-CZ
CZ.02.01.01/00/22_004/0007682



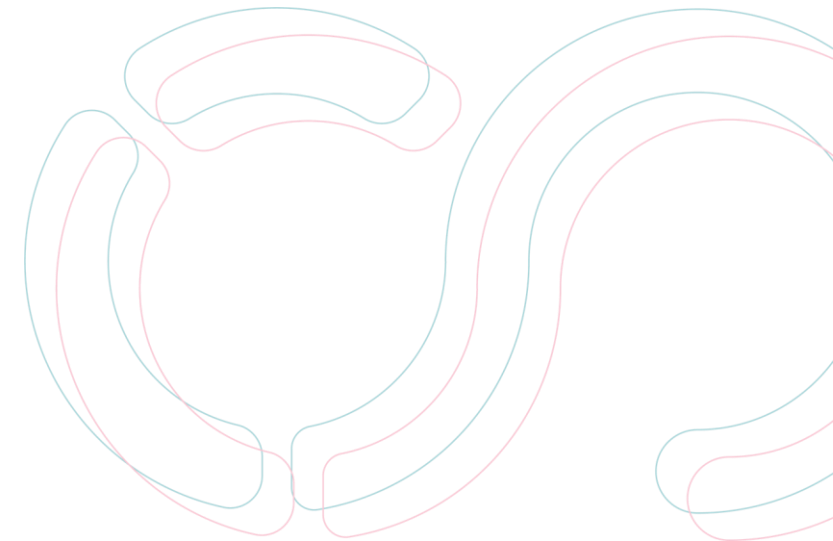
Definice

- ŘÍZENÝ SLOVNÍK JE:
- *Seznam slov nebo frází, které byly schváleny pro indexování. ([ČSN ISO 5127](#))*
- *Slovník lexikálních jednotek selekčního jazyka uspořádaný specifickým způsobem (např. zahrnuje vztahy ekvivalence, hierarchie a asociace), který slouží pro indexaci a vyhledávání dokumentů. ([TDKIV](#))*
- *Pevně definovaný soubor termínů, který usnadňuje organizaci a vyhledávání informací prostřednictvím standardizace a hierarchických vztahů mezi pojmy. (ChatGPT 3.5)*
- ***Slovník pořádacích znaků selekčního jazyka (systému organizace znalostí), který explicitně vyjadřuje vztahy mezi těmito znaky a určuje způsob jejich uspořádání; slouží pro popis, ukládání a vyhledávání informačních objektů. (JS)***

Řízené slovníky a lingvistika I.

JAZYK

- a) přirozený jazyk
- b) umělý jazyk
 - i) informační jazyk
 - (1) algoritmický informační jazyk
 - (2) logický informační jazyk
 - (a) databázový jazyk
 - (b) dotazovací jazyk
 - (3) strukturní informační jazyk
 - (a) značkovací jazyk
 - (4) selekční jazyk**
 - (a) identifikační SJ**
 - (b) věcný SJ**



Řízené slovníky a lingvistika II.

Přirozený jazyk

- univerzální
- asymetrický
- implicitní
- přirozený vývoj

selekční jazyk

- specializovaný
- symetrický
- explicitní
- umělý rozvoj

- *funkce a účel*
- *přesnost a jednoznačnost*
- *paradigmatické vztahy*
- *dynamičnost vs staticčnost*

Řízené slovníky a lingvistika III.

PARADIGMATICKÉ VZTAHY	Petr	jezdí	do školy	vlakem.	PJ
	student	cestování	vzdělávací instituce	dopravní prostředek	UM (SJ)
	osoba	pohyb	instituce	zařízení	
	subjekt	činnost	objekt	objekt	
	SYNTAGMATICKÉ VZTAHY				

Řízené slovníky a lingvistika IVa.

SLOŽKA

- *SLOVNÍK*
- *MORFOLOGIE*
- *SYNTAX*

PŘIROZENÝ JAZYK

- lexikum
- tvorba slov
- vztahy mezi slovy

SELEKČNÍ JAZYK

- řízený slovník
- konstrukce pořadacích znaků
- kombinace pořadacích znaků

Řízené slovníky a lingvistika IVb.

Selekční jazyk

- řízený slovník – pořádací znaky a vztahy mezi nimi
- konstrukce pořádacích znaků – pravidla pro tvorbu pořádacích znaků; např. inverze sousloví, binomická nomenklatura, základ a vlastnosti notace
- kombinace pořádacích znaků – pravidla pro spojování pořádacích znaků, tedy pravidla pro popis informačního objektu a pro formulaci rešeršního dotazu při vyhledávání

Řízené slovníky a lingvistika V.

Přirozený jazyk

Část nespokojených zemědělců společně s Asociací samostatných odborů ČR svolává na pondělí 19. února do Prahy protestní jízdu. Nelíbí se jim zemědělská politika, ale ani celá vláda a hrozí, že metropoli může blokovat víc než tisíc traktorů.

selekční jazyk A

zemědělec
odborová konfederace
Praha
stávka
národní zemědělská
politika
traktory

SELEKČNÍ JAZYK B

zemědělci
odborové svazy
Česká republika
stávky
zemědělská politika
kolové traktory

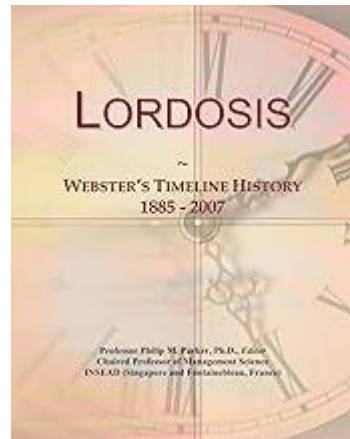
Základní vlastnosti ŘS

I. Primární účel ŘS

ŘS pro popis informačních objektů

klasifikace, číselníky, tezaury, předmětová
hesla, klíčová slova

Příklady: [MDT](#), [PSH](#), [LCSH](#), [AAT](#), [Agrovoc](#), [CZ-NACE](#)



Příklad použití:

[MeSH](#) vs [ICD](#)

[lordosis](#) vs [lordosis](#)

ŘS pro reprezentaci pojmů

taxonomie, ontologie

Příklady: [WordNet](#), [CRM-CIDOC](#), [OOPS](#), [DoCO](#)



Základní vlastnosti ŘS

II. Cílové atributy informačního objektu

IDENTIFIKAČNÍ (FORMÁLNÍ) ÚDAJE

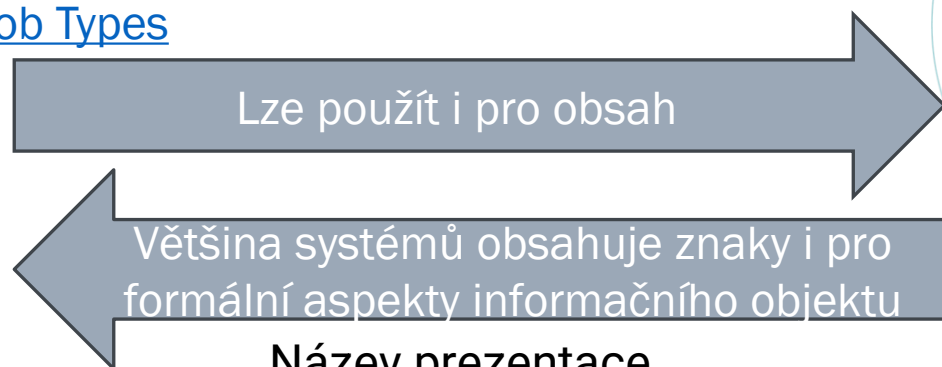
ŘS obsahuje pořadací znaky, které se vztahují k vnějším (formálním) aspektům informačního objektu, např. druh dokumentu, autor, vydavatel, místo vydání, jazyk atd.

Příklady: [Soubor jmenných autorit](#), [ISO 639 \(kódy jazyků\)](#), [GeoNames](#), [Music Job Types](#)

VĚCNÉ (OBSAHOVÉ, PŘEDMĚTOVÉ, TEMATICKÉ) ÚDAJE

ŘS obsahuje pořadací znaky, které se vztahují k vnitřním (obsahovým) aspektům informačního objektu.

Příklady: [MDT](#), [PSH](#), [LCSH](#), [AAT](#), [Agrovoc](#), [CZ-NACE](#)



Základní vlastnosti ŘS

III. Charakter řízeného slovníku

ŘS ENUMERATIVNÍHO CHARAKTERU

ŘS obsahuje výrazy pro všechny entity, které je potřeba popsat.

Typické pro: taxonomie, číselníky, tradiční předmětová hesla

ŘS SYNTETICKÉHO CHARAKTERU

Popis entit vzniká syntézou pořadacích znaků ŘS.

Typické pro: klíčová slova, tezaury, klasifikace

Základní vlastnosti ŘS

IV. Stupeň formalizace pořadacích znaků

ŘS ZALOŽENÉ NA PŘIROZENÉM JAZYKU

(PŘEDMĚTOVÉ SELEKČNÍ JAZYKY)

Lexikální jednotky jsou vyjádřeny jako jedno- nebo víceslovné výrazy převzaté ze slovníku přirozeného jazyka.

ŘS je uspořádán abecedně podle těchto lexikální jednotek, případné systematické uspořádání je sekundární.

Nejběžnější typy těchto ŘS: klíčová slova, předmětová hesla, tezaury, taxonomie

Příklady: [LCSH](#), [AAT](#), [Agrovoc](#), [PSH](#)

ŘS ZALOŽENÉ NA UMĚLÉM JAZYKU (KÓDU)

(SYSTEMATICKÉ SELEKČNÍ JAZYKY)

Klasifikační znaky jsou vyjádřeny notací (kódem), který je doplněn slovním vyjádřením (to ale není součástí klasifikačního znaku). Notační systém je obvykle vystaven na kategoriálním (logickém, systematickém) principu.

ŘS (klasifikační tabulky) je obvykle uspořádán systematicky podle principu, který je dán notačním systémem.

Nejběžnější typy těchto ŘS: klasifikace, číselníky, taxonomie, ontologie

Příklady: [MDT](#), [IPC](#), [ICS](#)

Základní vlastnosti ŘS

IV. Stupeň formalizace pořadacích znaků - příklady

ŘS ZALOŽENÉ NA PŘIROZENÉM JAZYKU

(PŘEDMĚTOVÉ SELEKČNÍ JAZYKY)

Typ I (základem lexikální jednotky je slovo nebo sousloví, resp. jednotlivý pojem):

stávky | zemědělská politika | ancient Italian pottery styles | Radschlepper | Emberiza leucocephalos | elektrický proud v kapalinách

Typ II (základem lexikální jednotky je složený výraz, resp. dva a více pojmů)

pokání – náboženské pojetí | Kordun (Chorvatsko) – dějiny – 19.-20. století | Cora language--Foreign elements--Spanish

ŘS ZALOŽENÉ NA UMĚLÉM JAZYKU (KÓDU)

(SYSTEMATICKÉ SELEKČNÍ JAZYKY)

Typ I (čistá notace – základem klasifikačního znaku je jeden typ znaků – obvykle písmena nebo číslice).

523.44 | RAT R | ATQL | 1061110002

Typ II (smíšená notace – základem klasifikačního znaku jsou různé typy znaků).

I55 | D03C 1/02 | J10.8 | |

Základní vlastnosti ŘS

IV. Stupeň formalizace pořadacích znaků – notační systém

- **Vybrané vlastnosti notačního systému**

- Počet pozic v jedné hierarchické úrovni

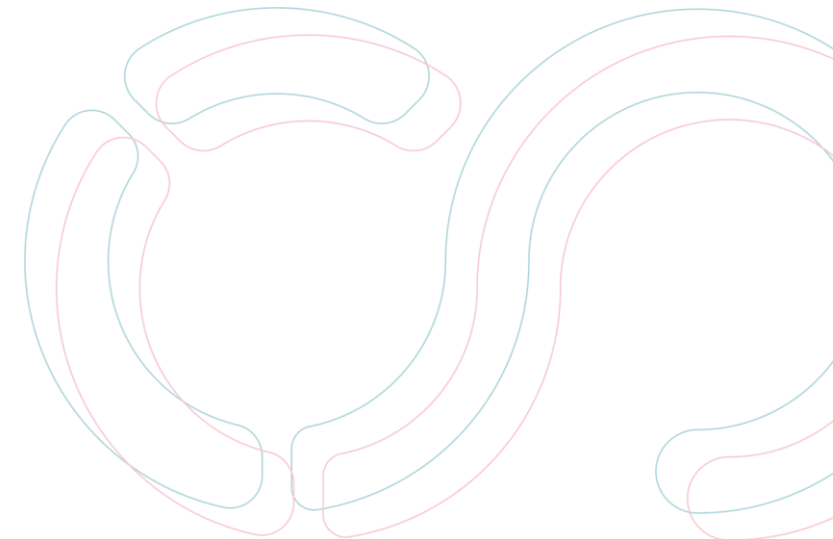
- Příklad: [MDT](#) x [ICS](#)

- Hierarchičnost/linearita

- Příklad: [MDT](#) x [BBC](#)

- Expanzivita notace

- Příklad: [MDT](#) x [LCC](#)



Základní vlastnosti ŘS

V. Vztahy mezi pořadacími znaky

ZÁKLADNÍ VZTAHY (SELEKČNÍ JAZYKY)

- hierarchie
 - rododruhový (generický) vztah
 - vztah celek-část (partitivní)
- asociace
 - vztah věcné souvislosti
- ekvivalence
 - vztah synonymie a kvazisynonymie

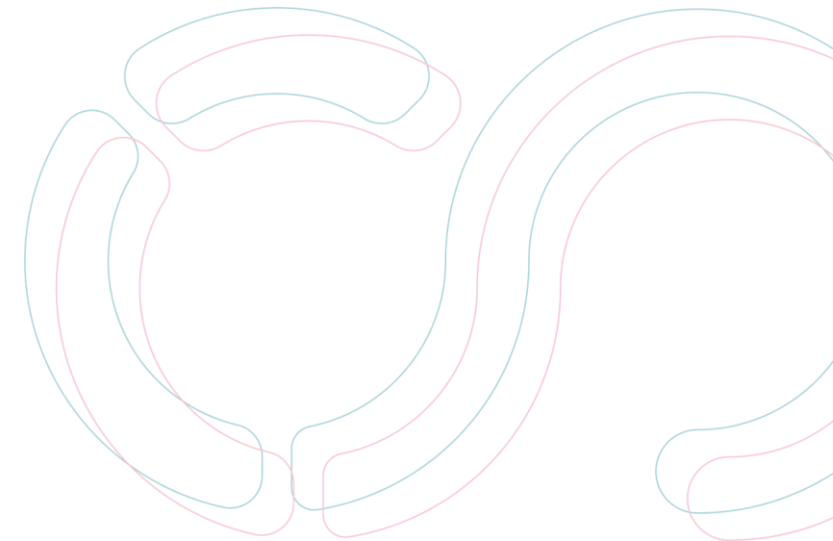
ROZVINUTÉ VZTAHY (ONTOLOGIE)

Libovolné typy vztahů

- P7 took place at (witnessed)
- P26 moved to (was destination of)
- P27 moved from (was origin of)
- P28 custody surrendered by (surrendered custody through)
- P37 assigned (was assigned by)
- P72 has language (is language of)
- P90 has value

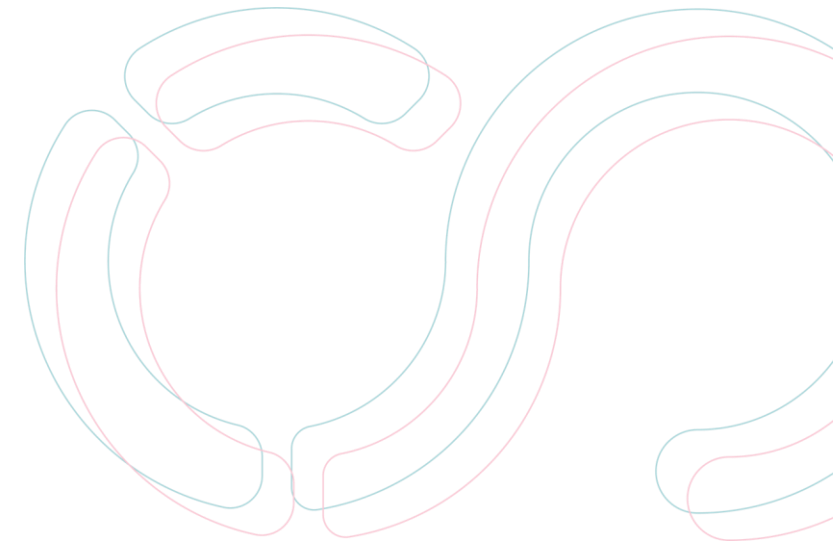
Standardizace vs lokalizace

- standardizace: na úrovni obecné struktury a pravidel
 - ČSN ISO 01 0193 (jednojazyčné tezaury)
 - ČSN ISO 01 0172 (vícejazyčné tezaury)
 - ČSN ISO 5964 (analýza dokumentů)
- harmonizace: na úrovni lexika
 - OKEČ → CZ-NACE
- lokalizace: na úrovni slovních vyjádření
 - klasikace
- lokalizace: na úrovni lexika
 - EET → ČPT



Popis informačních objektů

- gesce instituce
- procedura zpracování

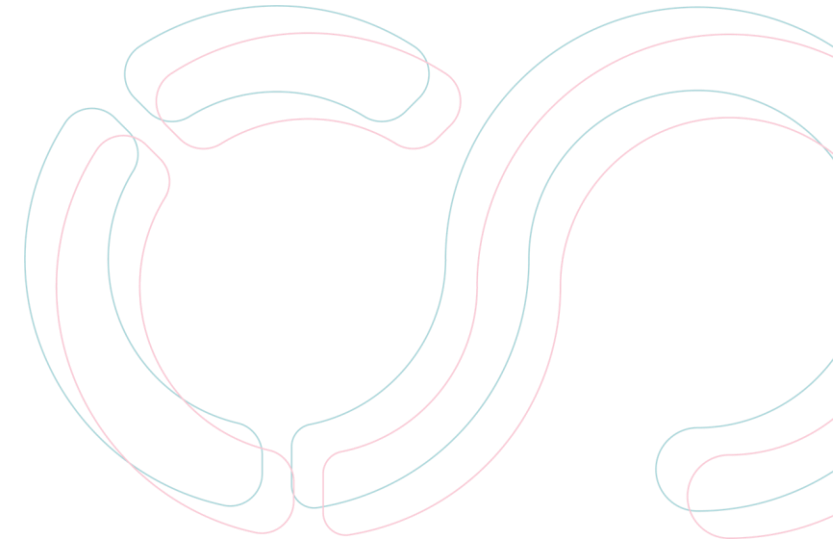


Praktické ukázky



Univerzální systémy

- Vybrané příklady:
 - LCSH
 - [portál](#)
 - MDT
 - [portál](#)



Oborové systémy

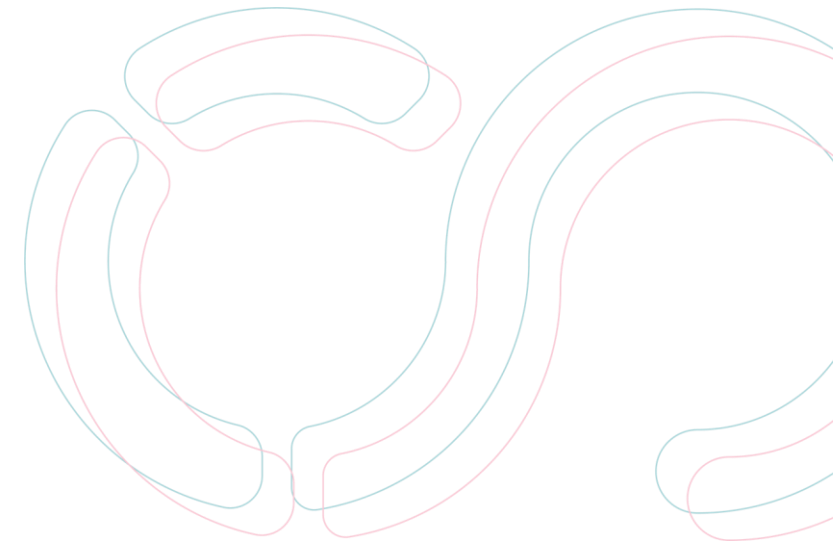
- Vybrané příklady:

- tezaury

- [AGROVOC](#)
- [Eurovoc](#)

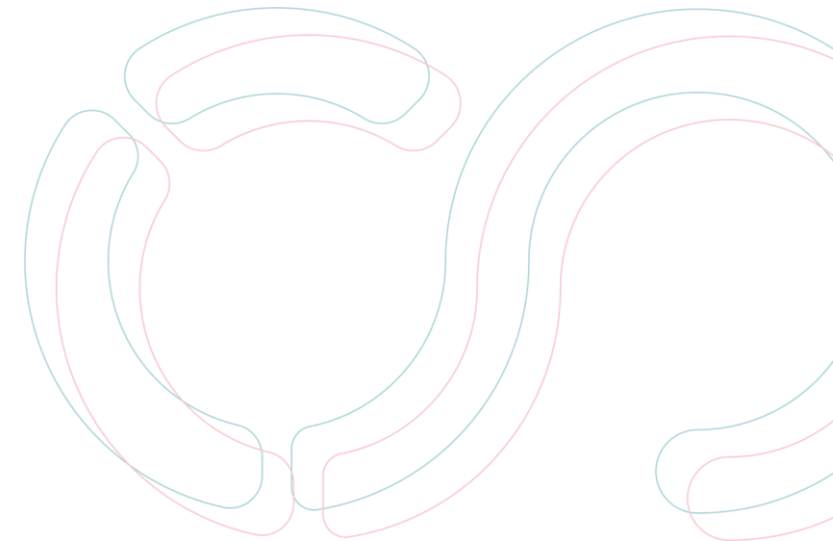
- klasifikace

- [MSC](#)
- [NACE + CZ-NACE](#)



Specializované systémy

- Vybrané příklady:
 - tezaury
 - [Backbone Thesaurus](#)
 - taxonomie
 - [TaDiRAH](#)
 - klasifikace
 - [ICS](#)
 - [IPC](#)
 - číselníky
 - [ISKN](#)



Děkuji za pozornost a zájem



Spolufinancováno
Evropskou unií



MUNI
ICS

cesnet

VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA

IT4INNOVATIONS
NÁRODNÍ SUPERPOČÍTAČOVÉ
CENTRUM

Registrační číslo IPs EOSC-CZ
CZ.02.01.01/00/22_004/0007682