



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

Metodika tvorby významových slovníků

Petr Křemen

Prerekvizita: Úvod do modelování významu dat ve veřejné správě



Co se dozvíte?

- jak identifikovat pojmy ve zdrojových datech
- jak vytvářet významové vazby mezi pojmy



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

Jak identifikovat pojmy?

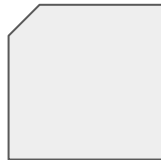


Odkud se pojmy berou?

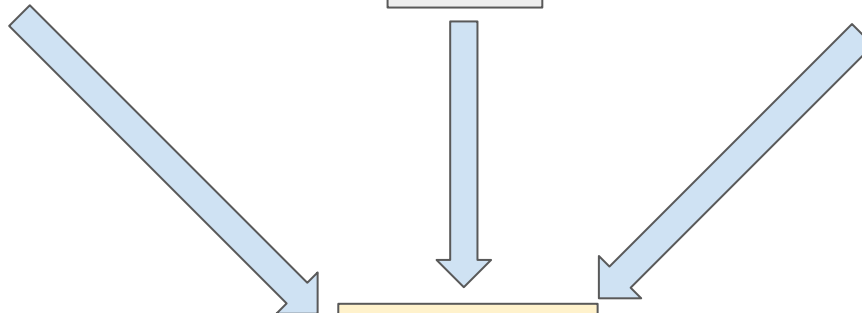
*Zákon, vyhláška,
evropská legislativa*

§

*Norma,
vnitroorganizační
předpis, metodika*



Expertní znalost



Pojem



Charakteristiky pojmu

název

- jednoznačný a přesný,
- v prvním pádě/první osobě jednotného čísla,
- začíná malým písmenem, nejedná-li se o vlastní jméno.

alternativní název

- zkratka/synonymum, používané pro označení pojmu.

definice

- pouze jedna,
- odkazuje na nadřazený pojem a omezující podmínky.

zdroj

- odkaz (nejlépe hypertextový) na nejmenší část dokumentu, ze které jednoznačně plyne význam pojmu (např. odkaz na definici).



(1) Legislativní pojem s definicí

Zákon č. 56/2001 Sb. *(Zákon o podmínkách provozu vozidel na pozemních komunikacích ...)*

...

§ 2

Základní pojmy

(1) Silniční vozidlo je motorové nebo nemotorové vozidlo, které je vyrobené za účelem provozu na pozemních komunikacích pro přepravu osob, zvířat nebo věcí



(1) Legislativní pojem s definicí

Zákon č. 56/2001 Sb. (Zákon o podmínkách provozu vozidel na pozemních komunikacích ...)

...

§ 2

Základní pojmy

(1) Silniční vozidlo je motorové nebo nemotorové vozidlo, které je vyrobené za účelem provozu na pozemních komunikacích pro přepravu osob, zvířat nebo věcí

Pojem

název: silniční vozidlo

definice: motorové nebo nemotorové vozidlo, které je vyrobené za účelem provozu na pozemních komunikacích pro přepravu osob, zvířat nebo věcí.

zdroj: <https://www.zakonyprolidi.cz/cs/2001-56#p2-1>



(2) Pojem ≠ odstavec zákona

Zákon č. 56/2001 Sb. *(Zákon o podmínkách provozu vozidel na pozemních komunikacích ...)*

...

§ 4

Registr silničních vozidel

...

(2) V registru silničních vozidel se u silničního vozidla uvádí

a) údaje o vlastníkovi a provozovateli, není-li totožný s vlastníkem, kterými jsou

1. jméno, popřípadě jména a příjmení, adresa místa trvalého pobytu, dlouhodobého pobytu, přechodného pobytu v délce alespoň 6 měsíců nebo jiného povoleného pobytu a rodné číslo, pokud bylo přiděleno, popřípadě datum narození,

Pojem:

- **název:** *jméno, popřípadě jména a příjmení, adresa místa trvalého pobytu, dlouhodobého pobytu, přechodného pobytu v délce alespoň 6 měsíců nebo jiného povoleného pobytu a rodné číslo, pokud bylo přiděleno, popřípadě datum narození,*
- **zdroj:** <https://www.zakonyprolidi.cz/cs/2001-56#p4-2-a-1>



(2) Pojem ≠ odstavec zákona

Zákon č. 56/2001 Sb. *(Zákon o podmínkách provozu vozidel na pozemních komunikacích ...)*

...

§ 4

Registr silničních vozidel

...

(2) V registru silničních vozidel se u silničního vozidla uvádí

a) údaje o vlastníkovi a provozovateli, není-li totožný s vlastníkem, kterými jsou

1. jméno, popřípadě jména a příjmení, adresa místa trvalého pobytu, dlouhodobého pobytu, přechodného pobytu v délce alespoň 6 měsíců nebo jiného povoleného pobytu a rodné číslo, pokud bylo přiděleno, popřípadě datum narození,

- Jak by vypadal konkrétní datový záznam takového údaje?
- Jak by bylo možné jej využít?
- Jak by bylo složité zkontrolovat jeho správnost?

Pojem:

- **název:** **jméno, popřípadě jména a příjmení, adresa místa trvalého pobytu, dlouhodobého pobytu, přechodného pobytu v délce alespoň 6 měsíců nebo jiného povoleného pobytu a rodné číslo, pokud bylo přiděleno, popřípadě datum narození,**
- **zdroj:** **<https://www.zakonyprolidi.cz/cs/2001-56#p4-2-a-1>**



(3) Pojem ≠ související sousloví

Zákon č. 56/2001 Sb. *(Zákon o podmínkách provozu vozidel na pozemních komunikacích ...)*

...

§ 4

Registr silničních vozidel

...

(2) V registru silničních vozidel se u silničního vozidla uvádí

a) údaje o vlastníkovi a provozovateli, není-li totožný s vlastníkem, kterými jsou

1. **jméno, popřípadě jména a příjmení,** adresa místa trvalého pobytu, dlouhodobého pobytu, přechodného pobytu v délce alespoň 6 měsíců nebo jiného povoleného pobytu a rodné číslo, pokud bylo přiděleno, popřípadě datum narození,

Pojem:

- **název:** **jméno, popřípadě jména a příjmení**
- **zdroj:** <https://www.zakonyprolidi.cz/cs/2001-56#p4-2-a-1>



(3) Pojem ≠ související sousloví

Zákon č. 56/2001 Sb. (Zákon o podmínkách provozu vozidel na pozemních komunikacích ...)

...

§ 4

Registr silničních vozidel

...

(2) V registru silničních vozidel se u silničního vozidla uvádí

a) údaje o vlastníkovi a provozovateli, není-li totožný s vlastníkem, kterými jsou

- jméno, popřípadě jména a příjmení, adresa místa trvalého pobytu, dlouhodobého pobytu, přechodného pobytu v délce alespoň 6 měsíců nebo jiného povoleného pobytu a rodné číslo, pokud bylo přiděleno, popřípadě datum narození

Člověk má "jméno", nebo více "jmen", ale nemá "jméno popřípadě jména a příjmení"?

Pojem:

- **název:** jméno, popřípadě jména a příjmení
- **zdroj:** <https://www.zakonyprolidi.cz/cs/2001-56#p4-2-a-1>



(3) Pojem ≠ související sousloví

Zákon č. 56/2001 Sb. *(Zákon o podmínkách provozu vozidel na pozemních komunikacích ...)*

...

§ 4

Registr silničních vozidel

...

(2) V registru silničních vozidel se u silničního vozidla uvádí

a) údaje o vlastníkovi a provozovateli, není-li totožný s vlastníkem, kterými jsou

1. **jméno** popřípadě jména a příjmení, adresa místa trvalého pobytu, dlouhodobého pobytu, přechodného pobytu v délce alespoň 6 měsíců nebo jiného povoleného pobytu a rodné číslo, pokud bylo přiděleno, popřípadě datum narození,

název pojmu - v případě, že "**jméno**" je v dokumentu použito ve více významech, je třeba pro účely identifikace pojmu název upřesnit.

Pojem:

- **název:**

- **alternativní název:**

- **zdroj:**

jméno vlastníka

jméno

<https://www.zakonyprolidi.cz/cs/2001-56#p4-2-a-1>



(4) Konsensuální pojem

Nenalezneme-li vhodný pojem v existujícím slovníku, zavádíme **nový pojem**. Jeho definici vytvoříme tak, aby:

- obsahovala obecnější pojem - **významového rodiče**
- co nejpřesněji popisovala **podmínky** a omezení, kterými se definovaný pojem liší od významového rodiče



(4) Konsensuální pojem

Nenalezneme-li vhodný pojem v existujícím slovníku, zavádíme **nový pojem**. Jeho definici vytvoříme tak, aby:

- obsahovala obecnější pojem - **významového rodiče**
- co nejpřesněji popisovala **podmínky** a omezení, kterými se definovaný pojem liší od významového rodiče

Pojem:

- *název:* **informační systém**
- *definice:* **systém vzájemně propojených prostředků a procesů, které slouží k ukládání, zpracování a poskytování informací.**
- *alternativní název:*
 - **IS**



(4) Konsensuální pojem

Nenalezneme-li vhodný pojem v existujícím slovníku (nebo nalezneme podobný s nevhodnou definicí), zavádíme **nový pojem**. Jeho definici vytvoříme tak, aby:

- vymezovala pojem vůči jinému obecnějšímu pojmu - **významovému rodiči**
- co nejpřesněji popisovala **podmínky** a omezení, kterými se definovaný pojem liší od významového rodiče

podmínky

Název
významového
rodiče

pojem:

název: **informační systém**

- *definice:* **system vzájemně propojených prostředků a procesů, které slouží k ukládání, zpracovávání a poskytování informací.**

- *alternativní název:*

- IS

Zkratka



Tezaurus - pojmová hierarchie pro snazší vyhledávání pojmů

Hierarchická vazba může mít různé významy. Vyhledáváme-li data o **silničních vozidlech**, může nám vyhledávač nabídnout i data související s **motory**, **osobními automobily**, nebo přímo s automobilem RZ 1AB 2345 - tedy s pojmy níže v hierarchii.

dopravní prostředek
silniční vozidlo

motor

automobil RZ 1AB 2345

osobní automobil

Část **silničního vozidla**

Konkrétní **vozidlo**

Speciální typ **vozidla**

Významový
rodič pojmu
**osobní
automobil**

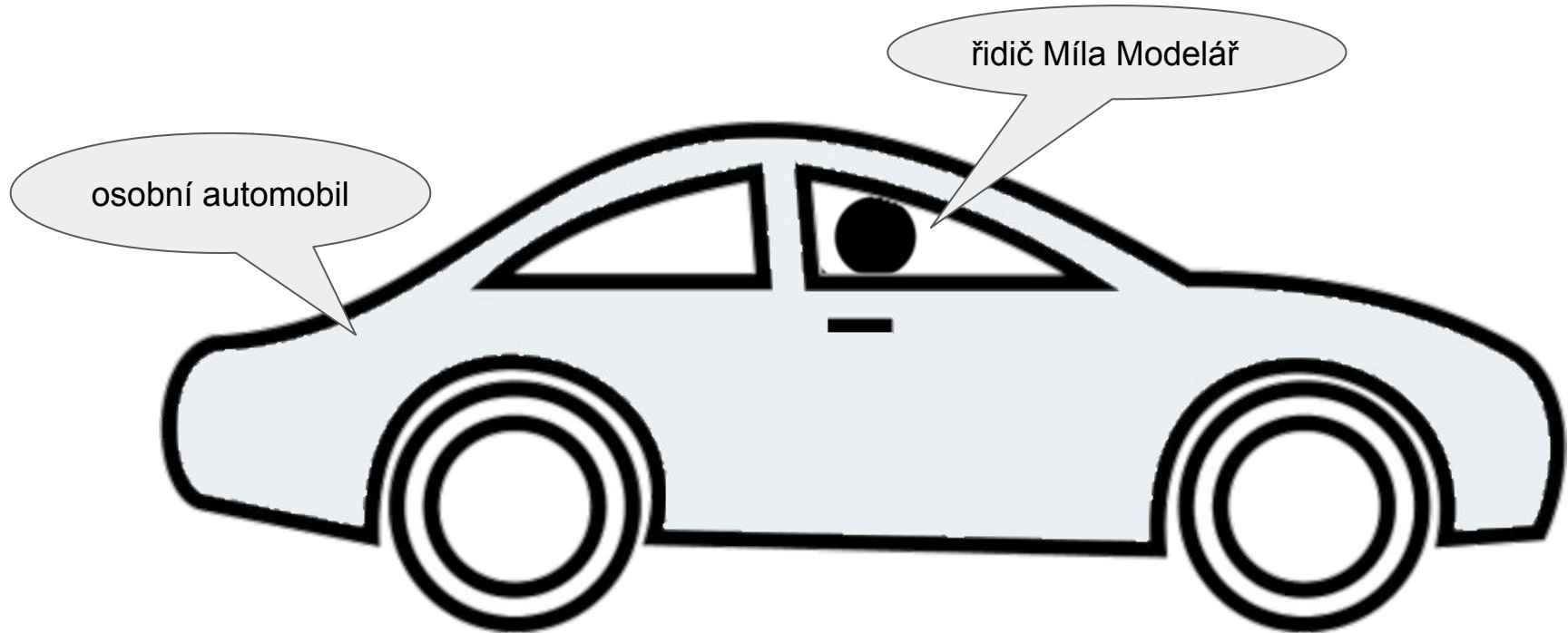


Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

Jak vytvářet významové vazby mezi pojmy?



Od reality k modelu





Pojem = základ konceptuálního modelu



osobní automobil

je oprávněn řídit



řidič Míla Modelář



Modelujeme



osobní automobil

je oprávněn řídit



řidič Míla Modelář

druh silničního vozidla

je oprávněn řídit

řidič



Kardinalitní omezení (1)



osobní automobil

je oprávněn řídit



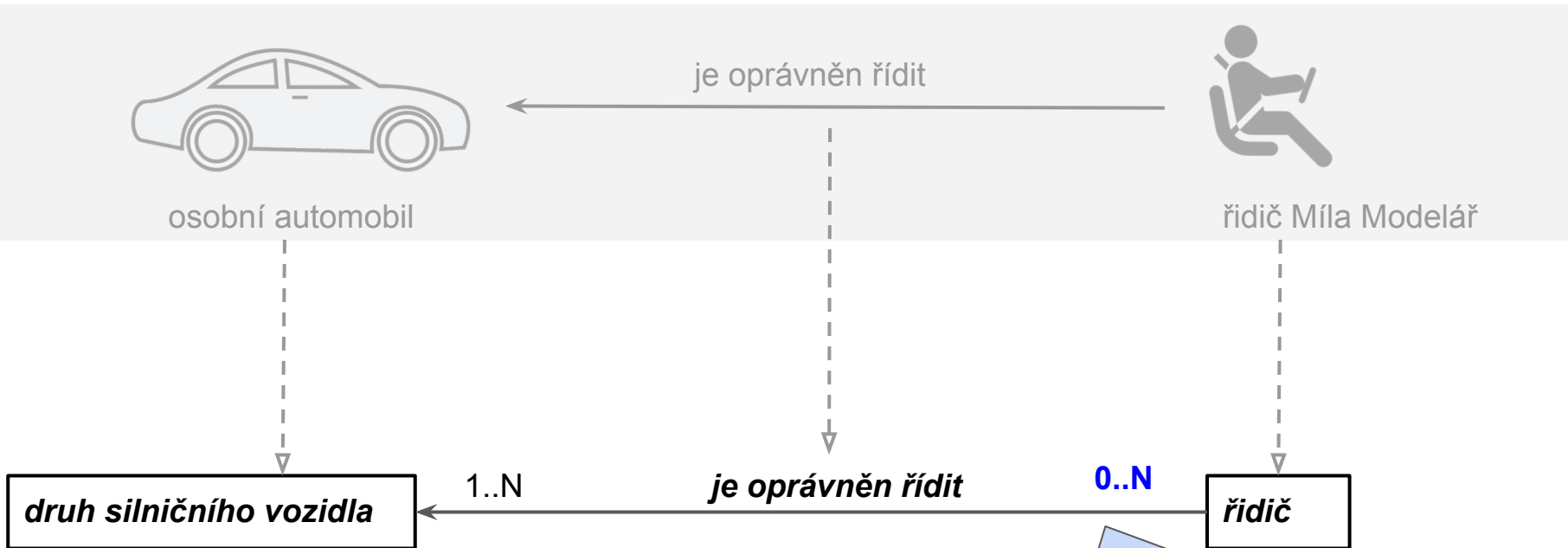
řidič Míla Modelář



1 = řidič je oprávněn řídit alespoň jeden druh silničního vozidla
N = řidič může mít více oprávnění k řízení různých druhů silničních vozidel



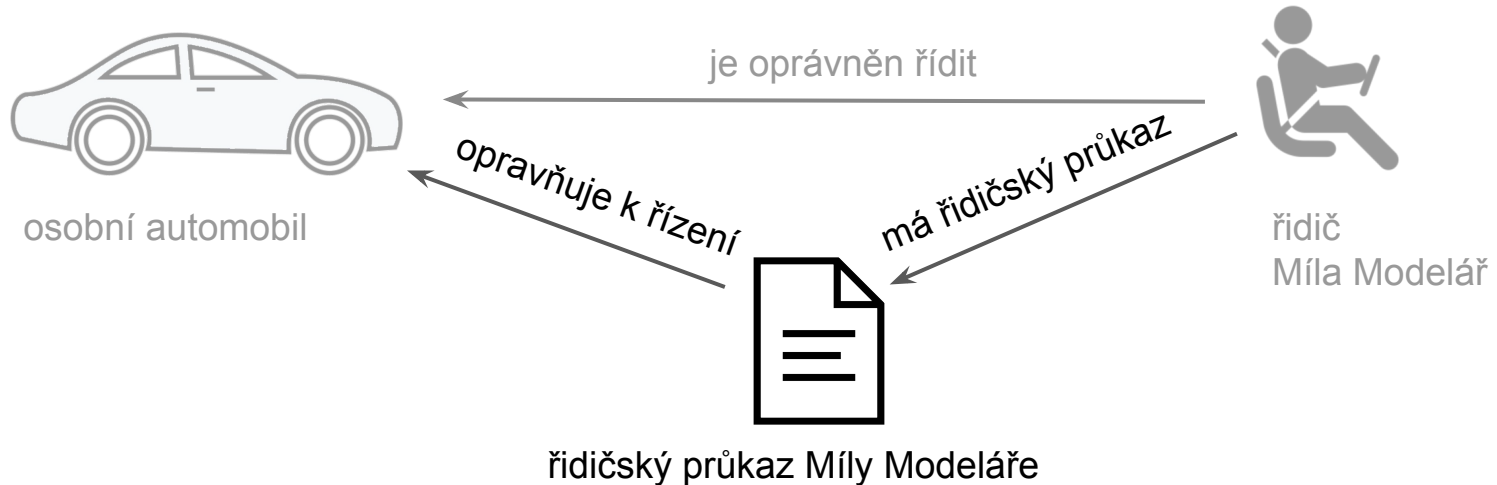
Kardinalitní omezení (2)



0 = pro daný druh silničního vozidla nemusí existovat žádné řidičské oprávnění
N = více řidičů může být oprávněno k řízení jednoho druhu silničního vozidla



Rozšiřujeme model (1)





Rozšiřujeme model (2)



osobní automobil

je oprávněn řídit



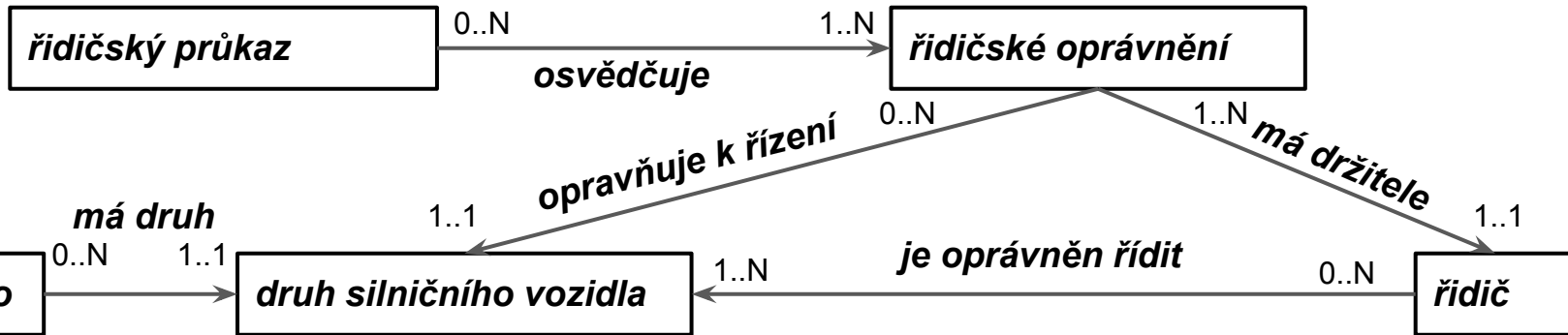
řidič
Míla Modelář

opravňuje k řízení



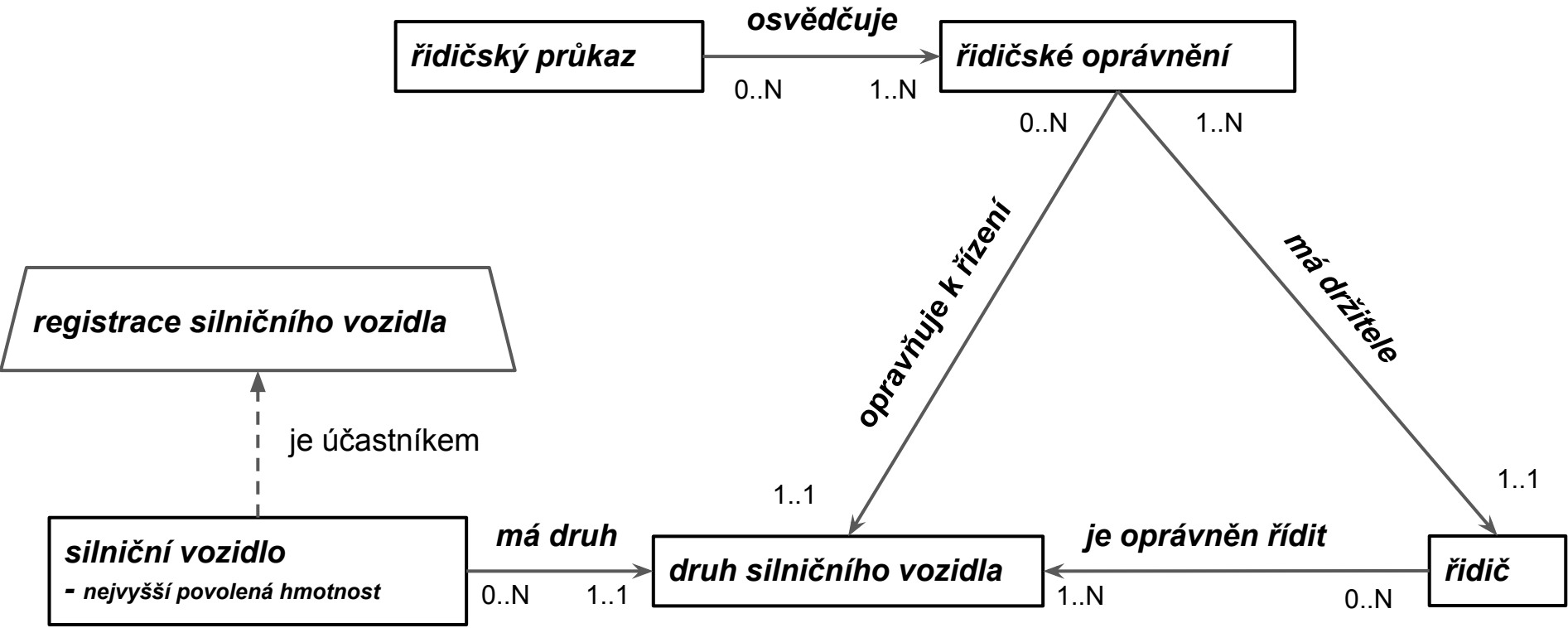
má řidičský průkaz

řidičský průkaz Míly Modeláře



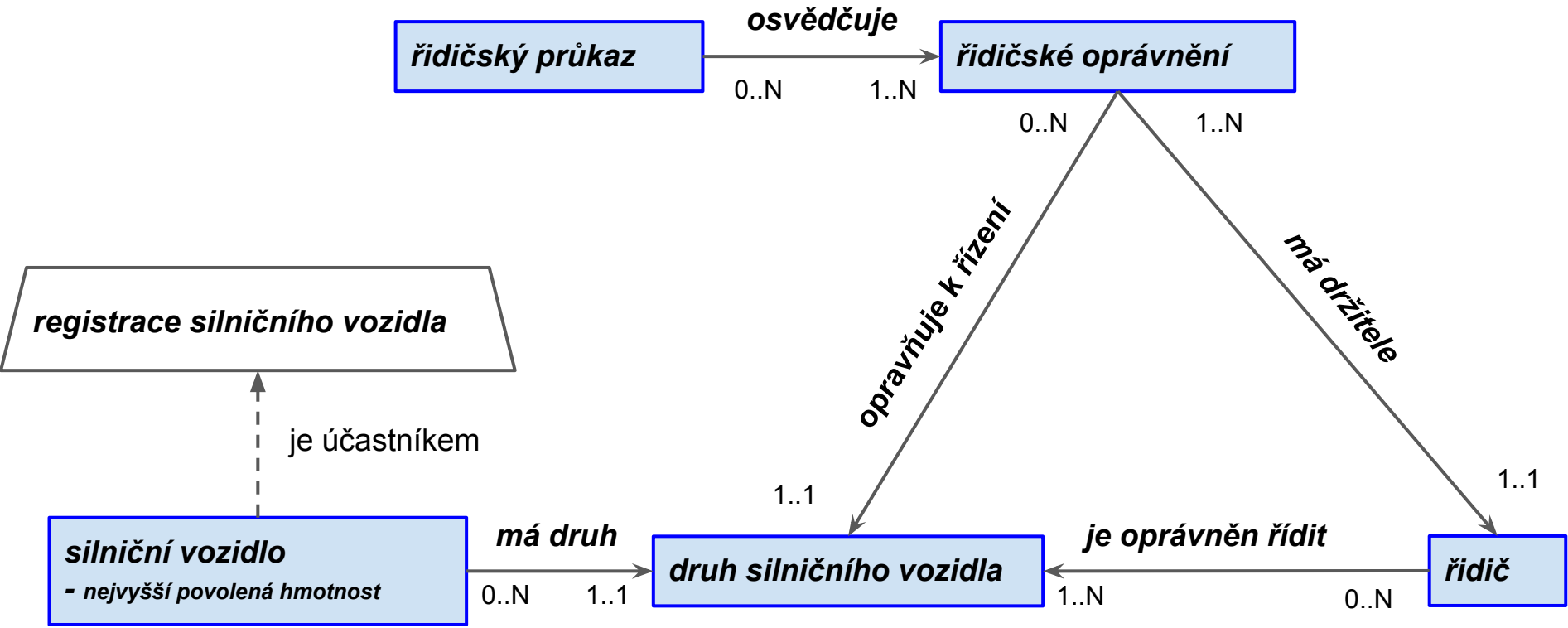


Rozšířený konceptuální model



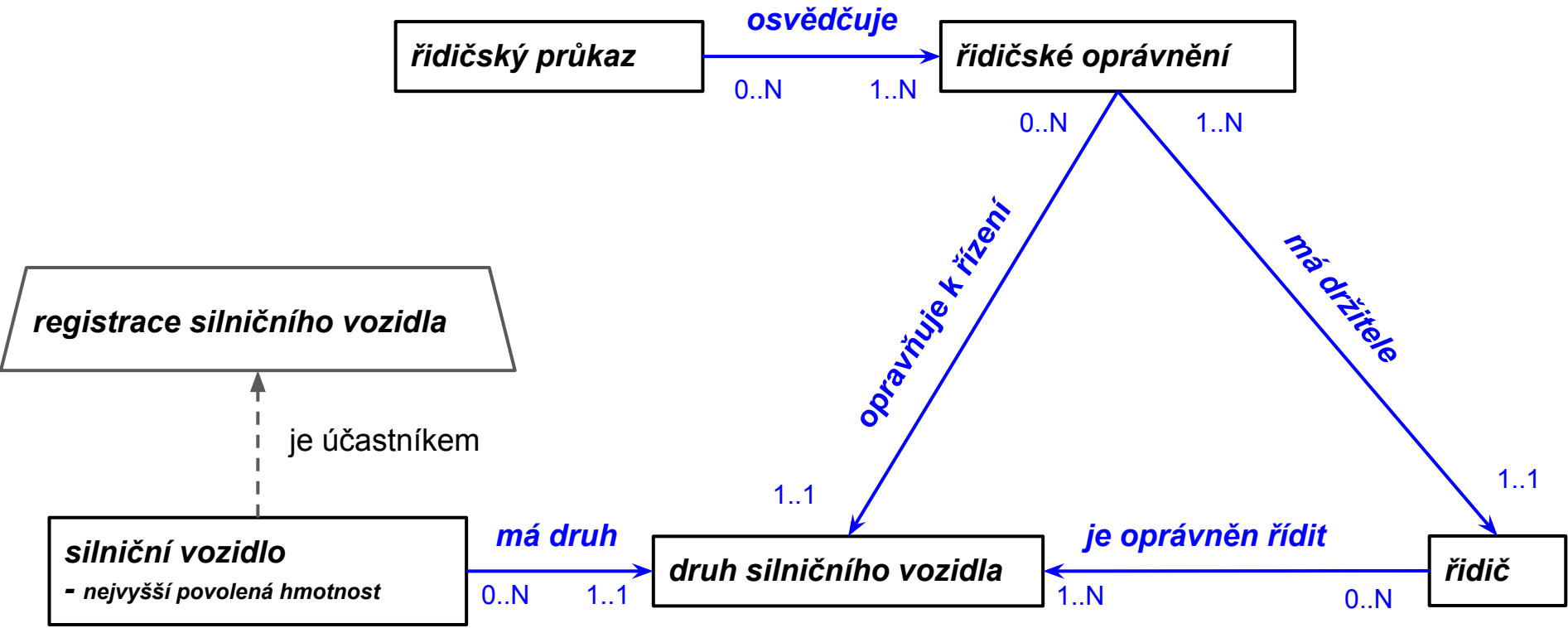


Typy objektů



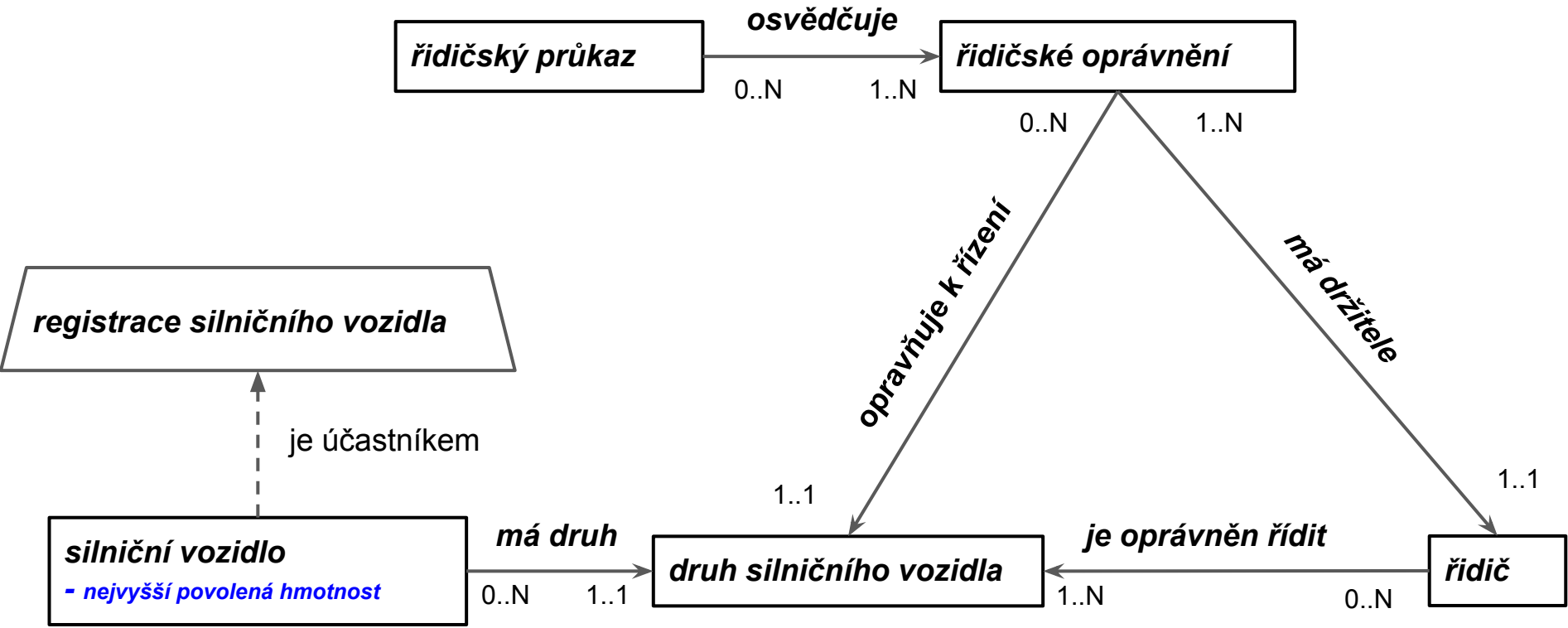


Typy vztahů



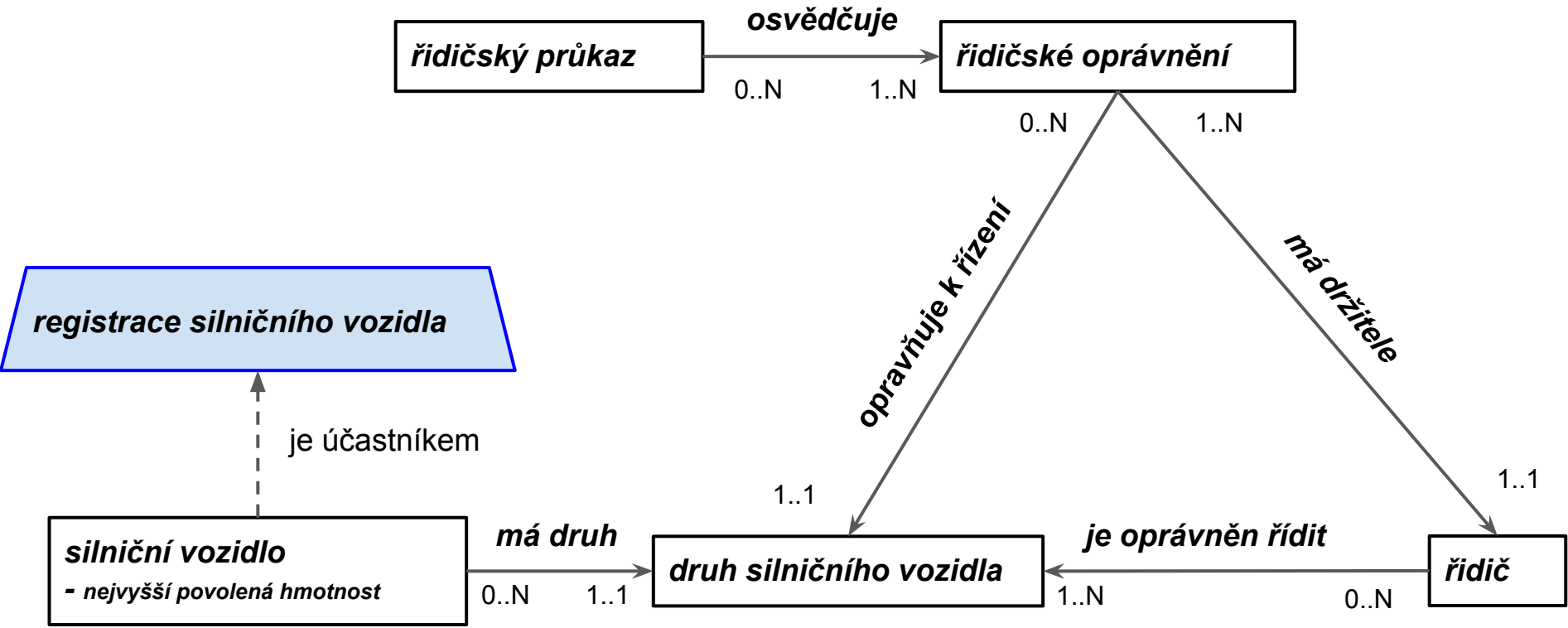


Typy vlastností





Typy událostí





Charakter pojmu?

- **Objekt** existuje sám o sobě
 - *silniční vozidlo, kolo, člověk, informační systém, ...*
- **Vlastnost** závisí na objektu
 - *hmotnost silničního vozidla, jméno fyzické osoby, ...*
- **Vztah** závisí na dvou nebo více objektech
 - *manželství, studium,*
- **Událost** se stala v čase
 - *svatba, přihlášení do kurzu, vydání řidičského průkazu*



Vhodné a nevhodné názvy pojmů

Za pojem pro účely tvorby sémantických slovníků považujeme **objekt, vlastnost, událost, vztah, nebo jejich typy**. Ne každé slovo či sousloví je vhodné jako název nějakého pojmu.

silniční vozidlo

silniční

nejvyšší povolená hmotnost silničního vozidla

nejvyšší povolená

registrace vozidla

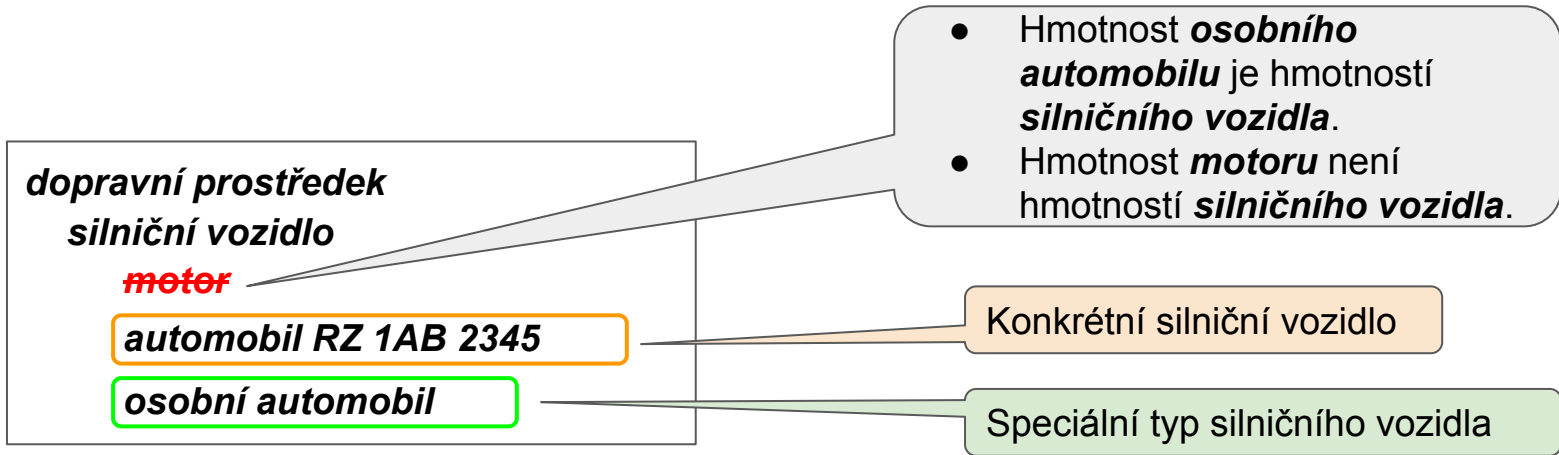
má držitele řidičského oprávnění





Jak přepoužít typy vlastností a vztahů?

Některé hierarchické vazby mohou sloužit k přepoužívání typů vlastností a vztahů.





Zpřesnění hierarchie pojmů pro kontrolu konzistence pojmů a přepoužívání vlastností

silniční vozidlo
- nejvyšší povolená hmotnost



osobní automobil

Osobní automobil má rovněž vlastnost ***nejvyšší povolená hmotnost***, ačkoliv jí v attributech nemusíme přímo uvádět - je zděděna od ***silničního vozidla***. Hodnota této vlastnosti samozřejmě může být pro každý druh silničního vozidla jiná.



Cheat Sheet

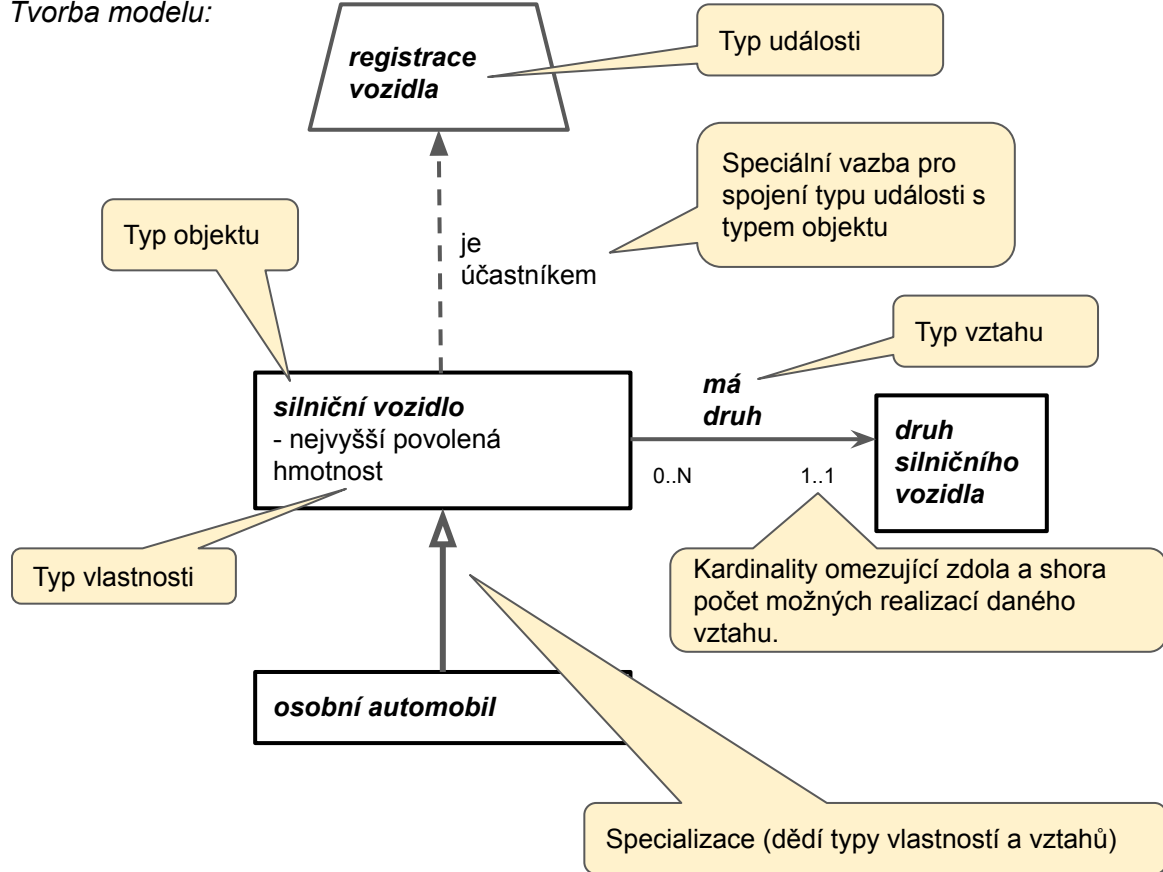
Postup tvorby sémantických slovníků

1. nalezneme pojmy, které již existují a lze je přepoužít
2. pokud vhodný pojem chybí, definujeme pojem nový
3. určíme charakter pojmů - objekt/vztah/vlastnost/událost, nebo jejich typ
4. propojíme pojmy pomocí významových vazeb

Tvorba pojmů:

- **název** - co nejjednoznačnější a nejpřesnější, který se uvádí v prvním pádě/první osobě jednotného čísla, začíná malým písmenem, nejedná-li se o vlastní jméno
- **alternativní název** - typicky zkratka/synonymum, která se běžně v určitém kontextu používá
- **definice** - nejvýše jedna
- **zdroj** - odkaz (nejlépe hypertextový) na co nejkonkrétnější část dokumentu (např. paragraf), kde je pojem definován, nebo ze kterého jednoznačně plyne jeho význam
- typicky označuje něco skutečného, z reálného světa

Tvorba modelu:





Další zdroje

- Detaily týkající se používaného modelovacího jazyka lze nalézt v [koncepti sémantického slovníku pojmů](#)
- Navazující moduly:
 - modul 3: Termlt
 - modul 4: OntoGrapher