

# ZNALOSTNÍ GRAFY JAKO NÁSTROJ PRO ANALÝZU SMART CITY

Viktor Beneš, Miroslav Svítek

Znalostní grafy, které poskytují platformu pro integraci a koordinaci různých aspektů městského plánování a správy, jsou zkoumány v kontextu hlavního města Prahy, konkrétně domény mobility. Článek představuje analýzu v oblasti korelace dat mezi emisemi, meteorologickou situací a automobilovou dopravou. Dále se zaměřuje na představení znalostních grafů jako nástroje pro vytváření holistických a integrovaných řešení pro výzvy související s městským rozvojem, pečlivého plánování a koordinaci při implementaci a využití znalostních grafů v praxi. Výsledky ukazují, že přestože znalostní grafy nabízejí mnoho možností, jejich úspěšné využití vyžaduje zohlednění specifického kontextu a potřeb každého města.

## Úvod

Znalostní grafy jsou v nejjednodušším případě reprezentovány jako soubor trojic, kde se každá trojice skládá ze subjektu, predikátu a objektu. Subjekt a objekt jsou entity, zatímco predikát reprezentuje vztah mezi nimi. Jednoduchý příklad na obr. 1 ukazuje vztah (predikát) autora (subjekt) ke svému článku (objekt).

Samotný znalostní graf znázorňuje provázanost jednotlivých atributů, které v případě města mohou být definovány jednotlivými městskými systémy včetně jejich detailů. Tento nástroj má potenciál dodat vyšší míru přehlednosti jednotlivým procesům ve městě včetně jednotlivých rolí, což může vést ke snazší digitalizaci procesů města a dále použití umělé inteligence k automatizaci jednotlivých městských procesů. Znalostní grafy hrají klíčovou roli v mnoha aplikacích umělé inteligence, včetně vyhledávání požadovaných informací, které budou strojově čitelné a významově (sémanticky) přesně definované. Znalostní grafy mohou být vytvořeny ručně nebo automaticky z textu v přirozeném jazyce [4] nebo z jiných strukturovaných dat.

Pro porozumění znalostním grafům je důležité si uvědomit hierarchii informačních agentů v rámci modelu zvaného

ho „DIKW Pyramid“ (Data-Information-Knowledge-Wisdom Pyramid), který ilustruje proces transformace surových dat (elementární informační konstrukt) na informace (vyšší úroveň významu, která k datům přidává další kontext, např. další datové sady apod.), které dále zvyšují svůj význam a stávají se z nich znalosti, jež následně vedou k získání moudrosti [1].

Znalostní graf tedy propojuje jednotlivé znalosti z určitých domén a dokáže zachytit i vysoký detail až na samotná elementární data.

Znalostní grafy lze klasifikovat do dvou hlavních kategorií: globální znalostní grafy a doménové znalostní grafy [2]. Globální znalostní grafy, například popis města, pokrývají široké spektrum obecných znalostí [3] a zahrnují celou řadu oblastí, které jsou většinou historicky dobře předdefinované. Na druhé straně, doménové znalostní grafy se zaměřují na specifickou oblast nebo doménu, jako je například mobilita, energetika, životní prostředí nebo odpadové hospodářství v rámci konceptu chytrých měst.

Samotné město, respektive pojem „Smart City“, je charakterizováno jako dynamický a neustále se vyvíjející komplexní adaptivní systém, který operuje v nejistém prostředí s mnoha zapojenými

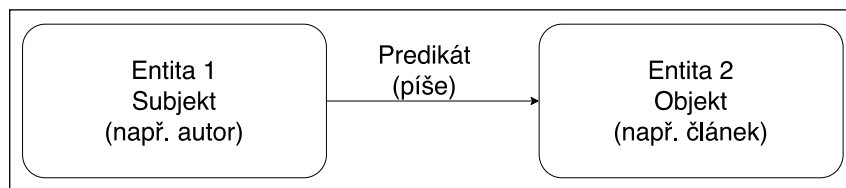
mi aktéry. Hlavním cílem tohoto systému je dosáhnout spokojenosti obyvatel a poskytnout jim pohodlné a bezpečné prostředí pro život [5].

Využití znalostních grafů v městském prostředí bylo popsáno v [6] včetně koordinace aktivit napříč různými městskými doménami, organizacemi a městskými částmi hl. m. Prahy. Tento příklad popisuje také městskou datovou platformu Golemio nejen jako nástroj pro agregaci a sdílení dat, ale i správu městských znalostí v kontextu jejich vzájemného propojení prostřednictvím znalostních grafů.

## Metodika

Tento článek navazuje na články [6] a [7] a předkládá ilustrativní příklad použití znalostního grafu pro potřeby dopravy, meteorologických podmínek a emisí.

Před implementací nástroje znalostních grafů je nezbytné zajistit sémantickou interoperabilitu, což znamená, že data musí být jednoznačně definovaná a srozumitelná napříč různými doménami. Příkladem může být definování termínu „vozidlo“ – odborníci z domény mobility jej mohou vnímat velice obecně jako dopravní prostředek. Oproti tomu například odborníci z domény odpadového hospodářství mohou termín „vozidlo“ vnímat jako specifické vozidlo pro odvoz odpadu. Je tedy třeba v první řadě sestavit dílčí slovníky pro jednotlivé domény včetně možných převodních tabulek, aby si odborníci z různých oblastí vzájemně rozuměli, ale aby si také v budoucnu rozuměly i systémy umělé inteligence napříč jednotlivými specializacemi.



Obr. 1: Znalostní graf – definice

## Výsledky

Ilustrativním příkladem budiž koordinace dopravy na základě dat o dopravě (intenzita dopravy a rychlost dopravního proudu), dat o meteorologických podmínkách (teplota vzduchu, akumulované srážky, směr a rychlost větru, nárazy větru) a lokální emise (CO, NO<sub>2</sub>, NO<sub>x</sub>, PM 10, PM 2.5). Tento příklad je vybrán v návaznosti na Klimatický plán hl. m. Prahy [8] řešící koordinaci dopravy v návaznosti na aktuální a budoucí meteorologické podmínky – jak se chovají cestující a emise za různých podmínek.

Článek [7] zobrazuje znalostní graf s identifikovanými vazbami mezi jednotlivými znalostními bázemi, konkrétně se jedná o téma dopravy, které zahrnuje dopravní situaci, O-D matici (*Origin-Destination Matrix*), která představuje zdroje a cíle v dopravě) a parkování, které je reprezentováno

P+R parkováním. Doprava zahrnuje také meteorologickou situaci a emise. Jedná se tak o propojení dvou domén – doménu mobility a doménu životního prostředí. Zároveň zde hraje svou roli i čas, který je zastoupen napříč doménami a vstupuje, respektive je součástí, všech znalostních bází.

Dopravní situace je ovlivněna aktuální O-D maticí, tedy odkud kam lidé cestují a jakým dopravním prostředkem. Tato O-D matice je však ovlivněna meteorologickou situací, dle které se každý jednotlivec rozhoduje. Tato meteorologická situace zároveň ovlivňuje i chování emisí v ovzduší, které jsou vytvářeny dopravní situací.

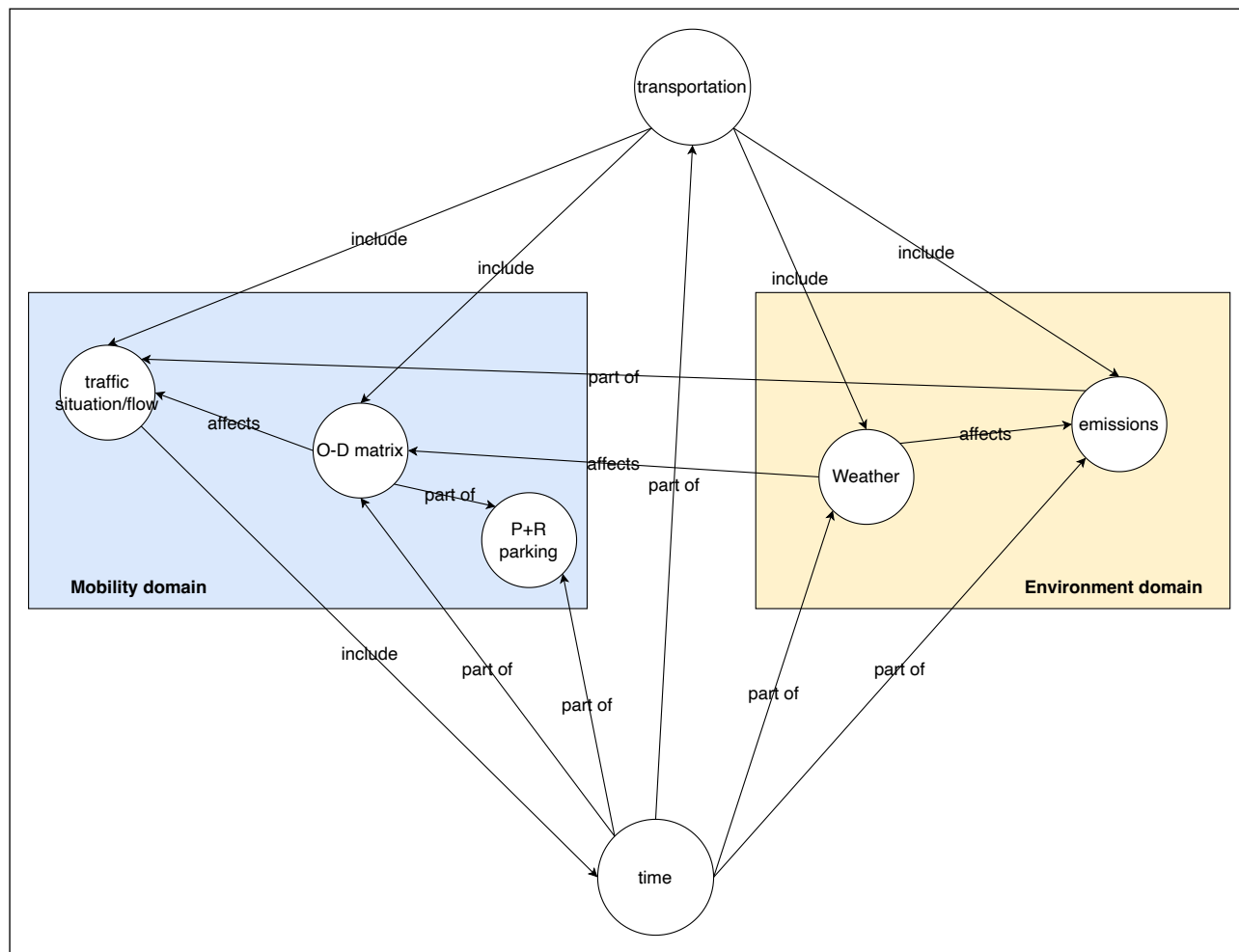
Jedná se tedy o jakýsi koloběh, kdy na sebe jednotlivé znalostní báze navazují, respektive se doplňují. Jednotlivé znalostní báze jsou ovšem rozdílné a mají svá specifika.

Analýza ověřovala existenci jednotlivých vztahů mezi znalostními bázemi na dvou vybraných lokalitách v hl. m. Praze prostřednictvím P-hodnoty a Pearsonova korelačního koeficientu R [9], který reprezentuje lineární regresi.

V tomto případě jsme porovnávali dopravní situaci – intenzitu dopravy a rychlost dopravního proudu s teplotou a akumulovanými srážkami a souhrnnou hodnotou CO<sub>2</sub> v daných lokalitách. Pro porovnání byly použity absolutní hodnoty jednotlivých veličin.

Obecný závěr analýzy ukazuje růst intenzity dopravy a rychlosti dopravního proudu s klesající teplotou. Je ovšem důležité zohlednit i samotnou situaci v dané lokalitě – parkování na silnici, respektive na krajnici, pruh vyhrazený pro MHD apod.

Dalším výsledkem analýzy je práce s absolutní hodnotou meteorologických



Obr. 2: Znalostní graf v oblasti dopravy [7]

1) Datová platforma hl. m. Prahy – Golemio. Dostupné z: <https://golemio.cz>. [vid. 10. 9. 2023].

údajů, která může sama o sobě přinést určitou volatilitu do výsledků. V těchto případech, kdy se jedná o rozhodování lidí na základě povětrnostních podmínek, je nevýznamný rozdíl například i 1 mm srážek. Do budoucna by se tedy mělo pracovat s kategorizovanými hodnotami pro srážky a minimální a maximální denní teplotou.

V původním přístupu nebyl zahrnut vítr, respektive směr a rychlost proudění vzduchu včetně jeho nárazů. Je tak třeba rozšířit spektrum sbíraných dat i o tyto proměnné a zkusit je opět kategorizovat, protože běžný člověk nemusí rozlišovat jednotlivé rychlosti větru, ale rozlišuje především jeho pocitovou sílu, která se odvíjí od absolutní rychlosti a může tak být kategorizována.

Výsledkem provedené analýzy je sada následujících doporučení pro další rozvoj:

- kategorizace meteorologických dat, alespoň do pocitových hodnot;
- přidání spektra atributů, zejména meteorologických hodnot, tj. rychlost a směr proudění vzduchu včetně větrných nárazů.

## Diskuse

Ukázkový případ demonstruje potenciál pro praktické využití nástroje znalostních grafů pro koordinaci dopravy v závislosti na předpovědi počasí, které umožní soulad s klimatickým plánem hl. m. Prahy.

Dle uvedených doporučení byl vypracován výpočet prostřednictvím P-hodnoty a Pearsonova korelačního koeficientu R. Tyto hodnoty reprezentují lineární regresi. Na základě uvedených doporučení byly výpočtem získány poznatky, které jednotlivé vztahy uvedené ve znalostním grafu výše potvrzují a dokáží stanovit váhu jednotlivých vazeb. Tato váha je ovšem omezena na výpočet, tedy na lineární regresi – je tedy třeba tyto vztahy dále zkoumat pro ověření regresního vztahu mezi znalostními bázemi.

Výsledky potvrzují silnou vazbu mezi dopravními atributy a hodnotami dusíku, tedy  $\text{NO}_2$  a  $\text{NO}_x$  v ovzduší. Tato vazba potvrzuje, že s rostoucí intenzitou dopravy roste i množství emisí, nicméně

s rostoucí rychlostí dopravního proudu jejich množství klesá. Lze z toho tedy odvodit následující: nižší intenzita dopravy, avšak rychlejší dopravní proud dokáže své lokálně produkované emise „vyvětrat“.

Zároveň je potvrzen silný vztah mezi rychlostí větru a emisemi  $\text{PM}_{10}$ ,  $\text{PM}_{2.5}$  a  $\text{CO}$ , kdy dochází k podobnému efektu a rychlé proudění větru tak dokáže tyto emise přesunout.

Tyto závěry mohou pomoci s digitalizací městských procesů. V tomto případě se může jednat o řízení dopravy, tedy integraci dat z různých zdrojů (např. kamery, průjezdové senzory, historické dopravní údaje apod.). Prostřednictvím znalostních grafů je možno předvídat dopravní kongesce ve větší míře, respektive na větší ploše řízeného území, a optimalizovat dopravní řízení. Znalostní graf kombinující různé datové zdroje (například údaje z průjezdových senzorů, meteorologické informace včetně předpovědi, data z navigačních systémů vozidel atd.) se správně nastavenými vazbami mezi datovými zdroji dokáže vytvořit komplexní přehled o pohybu vozidel, stanovení vzorců dopravních proudů včetně identifikace potenciálních bodů kongescí nejen v reálném čase, ale i do budoucnosti právě díky použití propojenosti dat a vzorců chování. Díky této kombinaci může řídicí systém okamžitě identifikovat a reagovat na různé dopravní situace, ať už jde o nehody, náhlé změny počasí nebo jiné nepředvídatelné události.

Optimalizace dopravy pomocí znalostních grafů pak spočívá v dynamickém přerozdělování dopravních proudů, například prostřednictvím světelného řízení či prostřednictvím kooperativních navigačních systémů. Taková adaptivní řešení zvyšují efektivitu celého dopravního systému a zároveň snižují negativní dopady dopravních kongescí na životní prostředí a kvalitu života obyvatel města.

Významnou výhodou znalostních grafů je přenositelnost postupu jejich vytváření na další subjekty (města), aby mohly sloužit jako univerzální nástroj pro standardizovanou koordinaci různých projektů, a to i napříč různými doménami města. Tato přenositelnost a schopnost kombinovat výsledky může

vést k vytvoření komplexních a integrovaných řešení, která zohledňují širokou škálu faktorů a všech zúčastněných stran.

Přístup použitý v tomto článku je níže pomocí SWOT analýzy porovnán s původním postupem uvedeným v [7]:

**Silné stránky:** Větší datová sada (rozšíření o datové sady); kategorické rozlišení meteorologických dat.

**Slabé stránky:** Více dat přináší větší náročnost na samotný sběr a předpracování dat a dále na výpočetní kapacitu.

**Příležitosti:** Databáze kategorií dat; stanovení hodnoty vazby mezi znalostními bázemi; přenositelnost na další města.

**Hrozby:** Nedostatek dat či jejich nezveřejnění.

## Závěr

Provedená analýza identifikovala otevřené body v předcházejících přístupech a vedla k návrhu úprav postupu, který zde byl představen.

Je užitečné porovnávat situace, které mají podobné parametry z důvodu správného nastavení. Srovnání obdobných situací může vést k lepšímu pochopení produkce lokálního znečištění a budoucího vývoje ve vztahu ke strategickým dokumentům města (Klimatický závazek [8]).

Dále se doporučuje provést analýzu s využitím většího souboru dat, tj. dat z více detektorů. Tyto informace pomohou určit řešení pro celou metropolitní oblast hl. m. Prahy. Provedení těchto analýz mezi všemi nejhoršími podmínkami nám pomůže postupně pochopit jednotlivé příčinné vícerozměrné vztahy.

Každá vazba ve znalostním grafu by měla být správně nastavena s příslušnou sémantikou, tedy slovníkem. Platforma znalostních grafů určitě pozitivně přispěje k holistickému řešení našich měst, protože umožní přenositelnost a efektivní spolupráci odborníků na jednotlivé domény a subdomény.

Nicméně, je důležité zdůraznit, že implementace a využití znalostních grafů v praxi vyžaduje pečlivé plánování a koordinaci. Každé město má svůj spe-

cifický kontext a potřeby, které je nutné zohlednit pro plné využití potenciálu znalostních grafů.

Znalostní grafy umožní například identifikovat prioritní opatření pro naplňování Klimatického plánu hl. m. města Prahy [8] a to ne ad-hoc opatřeními, ale opatřeními stanovenými na základě popsaných znalostních grafů, což umožní prioritizaci opatření, například dle Paretova pravidla.

#### Použité zdroje:

[1] VAN METER, H. J. (2020). Revising the DIKW Pyramid and the Real Relationship Between Data, Information, Knowledge and Wisdom. In: *Law, Technology and Humans* [on-line]. 2020, 2(2), 69-80. ISSN 2652-4074. [cit. 9. 7. 2023]. Dostupné z: doi:10.5204/lthj.1470.

[2] YANG, Y.-J., XU, B., HU, J.-W., TONG, M.-H., ZHANG, P. & ZHENG, L. (2018). Accurate and Efficient Method for Constructing Domain Knowledge Graph. In: *Ruan Jian Xue Bao / Journal of Software*. 29. 2931-2947. Dostupné z: doi:10.13328/j.cnki.jos.005552.

[3] SUCHANEK, F. M., KASNECI, G. & WEIKUM, G. (2007). Yago: a core of semantic knowledge. In: *Proceedings of the 16<sup>th</sup> international conference on World Wide Web* [on-line]. New York, NY, USA: ACM, 2007, s. 697-706. ISBN 9781595936547. [cit. 9. 7. 2023]. Dostupné z: doi:10.1145/1242572.1242667.

[4] KERTKEIDKACHORN, N. & ICHISE, R. (2018). An Automatic Knowledge Graph Creation Framework from Natural Language Text. In: *IEICE Transactions on Information and Systems* [on-line]. 2018, E101.D(1), 90-98. ISSN 0916-8532. [cit. 9. 7. 2023]. Dostupné z: doi:10.1587/transinf.2017SWP0006.

[5] SVÍTEK, M., SKOBELEV, P. & KOZHEVNIKOV, S. (2020). Smart City 5.0 as an Urban Ecosystem of Smart Services. In: BORANGIU, T., TRENTESAUX, D., LEITÃO, P., GIRET BOGGINO, A. & BOTTI, V. ed. *Service Oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future* [on-line]. Cham: Springer International Publishing, 2020, 2020-08-03, s. 426-438. Studies in Computational Intelligence. ISBN 978-3-030-27476-4. [cit. 9. 7. 2023]. Dostupné z: doi:10.1007/978-3-030-27477-1\_33.

[6] BENEŠ, V. & SVÍTEK, M. (2022). Knowledge graphs for Smart Cities. In: *2022 Smart City Symposium Prague (SCSP)* [on-line]. IEEE, 2022, s. 1-6. ISBN 978-1-6654-7923-3. [cit. 9. 7. 2023]. Dostupné z: doi:10.1109/SCSP54748.2022.9792541.

[7] BENEŠ, V. & SVÍTEK, M. (2023). Knowledge graphs for transport emissions concerning meteorological conditions. In: *2023 Smart City Symposium Prague (SCSP)* [on-line]. IEEE, 2023, s. 1-5. ISBN 979-8-3503-2162-3. [cit. 9. 7. 2023]. Dostupné z: doi:10.1109/SCSP58044.2023.10146219.

[8] MAGISTRÁT HLAVNÍHO MĚSTA PRAHY (2021). *Klimatický plán hlavního města Prahy do roku 2030. Praha na cestě k uhlíkové neutralitě.* [on-line]. Dostupné z: [https://klima.praha.eu/data/Dokumenty/Dokumenty%202023/klimaplan\\_cz\\_2301\\_09\\_online.pdf](https://klima.praha.eu/data/Dokumenty/Dokumenty%202023/klimaplan_cz_2301_09_online.pdf).

[9] BENEŠ, V. (2021). *Occupancy Prediction of Public Parking Spaces.* Diploma thesis, CTU in Prague, Faculty of Transportation Sciences. Dostupné z: <https://dspace.cvut.cz/handle/10467/94816>.

*Ing. Viktor Beneš, MSc.  
prof. Dr. Ing. Miroslav Svítek, dr. h. c.  
Ústav dopravní telematiky  
Fakulta dopravní  
ČVUT v Praze*

#### ENGLISH ABSTRACT

#### Knowledge Graphs as a Tool for Smart City Analysis, by Viktor Beneš & Miroslav Svítek

Knowledge graphs, which provide a platform for integrating and coordinating different aspects of urban planning and governance, are examined in the context of the capital city of Prague, specifically the mobility domains. The paper presents an analysis in the area of data correlation between emissions, meteorological situation and road traffic. It also focuses on introducing knowledge graphs as a tool for creating holistic and integrated approaches to urban development challenges, careful planning and coordination in the implementation and use of knowledge graphs in practice. The results show that although knowledge graphs offer many possibilities, their successful use requires taking into account the specific context and needs of each city.