



Data v praxi

Mgr. Jiří Fišer, Ph.D.



Ústí nad Labem 2021

- Předmět:** Data v praxi
- Studijní program:** Aplikovaná informatika
- Anotace:** Předmět formou exkurzí a seminářů vedených odborníky z praxe seznamuje studenty s různými zdroji dat, jejich aktuálním využitím a především potenciálem, které nabízí využití moderních přístupů zpracování a interpretace dat.

Jazyková korektura nebyla provedena, za jazykovou stránku odpovídá autor.

Obsah předmětu

Organizace studia v předmětu *Data v praxi* je shodná pro prezenční a kombinovanou formu studia.

Předmět seznamuje studenty s různými zdroji dat, jejich aktuálním využitím a především potenciálem, které nabízí využití moderních přístupů zpracování a interpretace dat.

Předmět využívá dvě základní formy výuky:

1. **exkurze ve firmách a organizacích**, které ve své praxi získávají nebo zpracovávají data (součástí exkurzí jsou prezentace jejich odborníků in situ),
2. **semináře vedené odborníky z obdobně orientovaných firem/organizací na katedře informatiky.**

Exkurze i semináře jsou pro studenty v kombinované formě studia povinné ve stejném rozsahu jako pro studenty prezenčního formě studia.

Obsah exkurzí a seminářů je vždy přizpůsoben aktuálním výzvám a stavu IT oddělení v příslušných firmách, a proto nelze obsah předmětu pokrýt statickým učebním textem. Semináře na půdě katedry jsou nahrávány a odkazy jsou k dispozici na Moodle serveru kacer.ujep.cz.

Prezentace odborníků z praxe (jak v rámci exkurzí, tak seminářů) se zaměřují na jednotlivé aspekty procesního řetězce (např. využití senzorických sítí, agregace dat, čištění a filtrování, ukládání a zabezpečení, vizualizace a interpretace). Konkrétně může jít například o problematiku zadání projektu na analýzu dat a s ním spojené právní a jiné úkony, sběr a získávání dat (šetření, open data aj.), oběh dat ve firmě a bezpečnost, jejich následná analýza a zpětné reportování. V rámci exkurze se studenti mohou seznámit s daným firemním prostředím a s příslušným technickým řešením (např. sběru dat) na místě. Kromě IT pracovníků se na výuce mohou podílet i odborníci, kteří data využívají a interpretují (manažeři, lékaři apod.).

Rozsah: tři exkurze každá v rozsahu 3 výukových hodin, tři semináře každý v rozsahu 2 výukových hodin

Přehled exkurzí a seminářů:

- *ADLER Czech, a.s.* (exkurze) – zpracování a vizualizace obchodních dat
- *BM COM s. r. o.* (exkurze) – IoT řešení (od sběru dat typicky ze senzorů s využitím sítí, jako je například LoRa, SigFox nebo NB-IoT, po jejich následné zpracování a analýzu s využitím platformy ThingWorx), vzdálený dohled CPE (routerů) v síti operátora (jejich vizualizace a využití z pohledu technické podpory).
- *Krajská zdravotní, a.s.* (seminář) – správa klinických a administrativních dat poskytovatele zdravotní péče
- *Process Automation Solutions s.r.o.* (exkurze) – využití dat v procesech automatizace

Dva semináře budou zajištěny dle aktuální nabídky firem resp. orgánů veřejné správy.

Seminární práce

Zápočet je v tomto předmětu udělován za vypracování seminární práce, která je primárně určena pro prezentující firmu či organizaci. Seminární práce má podobu dokumentu, který:

- **prezentuje alternativní zdroje dat využitelné v rámci dané firmy či organizace (včetně výhod, které přinášejí)**
- **určuje alternativní metody zpracování dat resp. alternativní výstupy tohoto zpracování (popis by měl být relevantní a dostatečně konkrétní pro účely daného poskytovatele dat)**
- **uvádí alternativní technologie zpracování dat se specifikací výhod oproti současnému řešení**
- **obsahuje jiná relevantní doporučení týkající se procesu zpracování dat**

Hlavním zdrojem dat pro seminární práci by měla být **řešerše existujících řešení** z výzkumných zdrojů a prezentací na internetu. Nezbytnou součástí dokumentu jsou proto **odkazy na použité zdroje** se standardními bibliografickými záznamy. Práce, která využívá neodkazované zdroje, nemůže být uznána.

Seminární práci hodnotí pracovník prezentující firmy, přičemž kritériem není rozsah, ale:

- **potenciální přínos pro danou firmu či organizaci**
- **konkrétnost a aplikovatelnost řešení**
- **počet a kvalita použitých zdrojů** (zdrojem nesmí být jiné seminární práce)

Formální požadavky na práci:

- minimální rozsah 5 normostran
- dodržení normy ČSN ISO 690:2011 pro bibliografické citace
- formát stránky A5 (pro snadnější čtení na přenosných zařízeních)
- formát PDF

Doporučené internetové zdroje

Před absolvováním předmětu resp. v jeho rámci lze využívat internetové zdroje **případů užití** (velkých) dat v různých oblastech lidské činnosti. Doporučit lze především následující odkazy (stav k začátku roku 2021).

Obecné případy užití

- 40 Stats and Real-Life Examples of How Companies Use Big Data
- Top 6 Data Science Use Cases that are Changing the World
- Hadoop Examples: 5 Real-World Use Cases

Zdroje případů užití podle oblastí

Analýza (ekonomických) dat

Top 9 Data Science Use Cases in Analytics

Telekomunikace

Top 10 Data Science Use Cases in Telecom

Zdravotnictví

14 Potential Use Cases of Data Science in Healthcare

Školství a akademická sféra

5 Ways Data Science Is Improving Higher Education

Doprava a logistika:

Applying Data Science Techniques to the Logistics Industry

Bankovníctví

4 Powerful Use Cases for Data Science in Finance

Přehled technologií (z pohledu využitelnosti)

MIT Technology Review (sekce data science a big data)

<https://www.technologyreview.com/tag/data-science/>

<https://www.technologyreview.com/hub/big-data-big-business>

Dostupné jsou i přehledy aktuálně využívaných technologií (i když následující přehled ukazuje, že některé technologie jsou již delší dobu aktuální).

Top 10 Big Data Technologies