

Juhaňák, Libor

## Vznik a historické kořeny analytiky a data miningu ve vzdělávání

In: Juhaňák, Libor. *Analytika učení a data mining ve vzdělávání v kontextu systémů pro řízení výuky*. Vydání první Brno: Masarykova univerzita, 2023, pp. 13-24

ISBN 978-80-280-0184-1; ISBN 978-80-280-0185-8 (online ; pdf)

Stable URL (handle): <https://hdl.handle.net/11222.digilib/digilib.77690>

Access Date: 29. 11. 2024

Version: 20230228

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

## 2 VZNIK A HISTORICKÉ KOŘENY ANALYTIKY A DATA MININGU VE VZDĚLÁVÁNÍ

Zaměříme-li se nejprve na výzkumnou oblast označovanou jako data mining ve vzdělávání (*educational data mining – EDM*), lze říci, že o vzniku této oblasti se nejčastěji hovoří ve spojení s lety 2008 a 2009. Je to dáno tím, že v roce 2008 proběhla první mezinárodní konference zaměřená specificky na data mining ve vzdělávání (*International Conference on Educational Data Mining*) a v roce 2009 pak vyšlo první číslo časopisu *Journal of Educational Data Mining*, zaměřeného výhradně na tuto oblast. Jak ovšem naznačuje jeden z prvních článků publikovaných ve zmiňovaném časopise (Baker & Yacef, 2009), historické kořeny této oblasti jsou výrazně starší. Pokud bychom vyšli ze studie Romera a Ventury (2007), pak můžeme spatřovat první snahy o využití data miningu ve vzdělávání již od roku 1995<sup>2</sup>. V kontextu e-learningu a technologiemi podporovaného učení se však začala data miningu věnovat větší pozornost spíše až na přelomu tisíciletí. To se následně projevilo i v tom, že se téma data miningu ve vzdělávání začalo postupně objevovat v podobě specializovaných workshopů a seminářů na několika tematicky příbuzných konferencích<sup>3</sup> (Bakhshinategh, Zaiane, ElAtia, & Ipperciel, 2018). Rostoucí zájem o téma data miningu v kontextu vzdělávání pak v roce 2005 vyústil v publikaci *Data Mining in E-learning* (2005), která se tak stala první knihou věnující se tomuto specifickému tématu. Zároveň bylo v roce 2005 poprvé použito označení

---

2 Zde však šlo ještě o využití data miningu v kontextu „tradičního“ vzdělávání nikoli v souvislosti s online vzdělávacími systémy typu LMS či VLE, jelikož ty se začaly více rozšiřovat až o několik let později.

3 Šlo např. o konference *International Conference on Artificial Intelligence in Education* (AIED), *International Conference on Intelligent Tutoring Systems* (ITS) či *Annual Conference of the Association for the Advancement of Artificial Intelligence* (AAAI).

## 2 Vznik a historické kořeny analytiky a data miningu ve vzdělávání

*Educational data mining* přímo v názvu odborného workshopu zaměřujícího se na tuto oblast (Romero & Ventura, 2013).

Přesto lze právě v souvislosti s lety 2008 a 2009 hovořit o určitém počátku, a to v tom smyslu, že přibližně v této době se začíná o data miningu ve vzdělávání hovořit jako o samostatné a svébytné výzkumné oblasti. Vedle již zmiňované první konference a prvního časopisu věnujících se tomuto tématu lze za další důležitý milník považovat knihu *Handbook of Educational Data Mining* (Romero, Ventura, Pechenizkiy, & Baker, 2010), která se jako první zaměřuje explicitně na oblast EDM. V následujícím roce pak byla založena mezinárodní společnost pro data mining ve vzdělávání (*The International Educational Data Mining Society*). Zároveň se tou dobou začínala vyčleňovat druhá z oblastí, jimž se věnuje tato publikace, totiž analytika učení (*learning analytics – LA*). V souvislosti s oběma zmiňovanými oblastmi se pak přibližně od roku 2014 postupně začíná objevovat hned několik odborných knih, které jsou pro přehlednost shrnuty v následující tabulce (viz tabulka 1).

**Tabulka 1:** Přehled základních odborných knih věnujících se specificky problematice data miningu ve vzdělávání či analytice učení

Rok	Oblast	Název knihy	Odkaz na zdroj
2005	EDM	<i>Data Mining in E-learning</i>	(Romero & Ventura, 2005)
2010	EDM	<i>Handbook of Educational Data Mining</i>	(Romero, Ventura, Pechenizkiy, & Baker, 2010)
2014	EDM	<i>Educational Data Mining: Applications and Trends</i>	(Peña-Ayala, 2014a)
2014	LA	<i>Learning Analytics: From Research to Practice</i>	(Larusson & White, 2014)
2016	LA	<i>Developing Effective Educational Experiences through Learning Analytics</i>	(Anderson & Gavan, 2016)
2016	EDM i LA	<i>Data Mining and Learning Analytics: Applications in Educational Research</i>	(ElAtia, Ipperciel, & Zaiane, 2016)
2016	EDM	<i>Educational Data Mining with R and Rattle</i>	(Kamath & Kamat, 2016)
2017	LA	<i>Big Data and Learning Analytics in Higher Education: Current Theory and Practice</i>	(Daniel, 2017)
2017	EDM i LA	<i>Big Data in Education: The digital future of learning, policy and practice</i>	(Williamson, 2017)
2017	LA	<i>Handbook of Learning Analytics</i>	(Lang, Siemens, Wise, & Gašević, 2017)
2017	LA	<i>Learning Analytics: Fundamentals, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning</i>	(Peña-Ayala, 2017)
2017	LA	<i>Learning Analytics Explained</i>	(Sclater, 2017)

## 2 Vznik a historické kořeny analytiky a data miningu ve vzdělávání

Rok	Oblast	Název knihy	Odkaz na zdroj
2017	LA	<i>Learning and Knowledge Analytics in Open Education</i>	(Lai & Lehman, 2017)
2018	LA	<i>Learning Analytics Goes to School: A Collaborative Approach to Improving Education</i>	(Krumm, Means, & Bienkowski, 2018)
2018	LA	<i>Learning Analytics in the Classroom: Translating Learning Analytics Research for Teachers</i>	(Lodge, Horvath, & Corrin, 2018)
2018	LA	<i>Learning Analytics in Higher Education: Current Innovations, Future Potential, and Practical Applications</i>	(Lester, Klein, Johri, & Rangwala, 2018)
2019	LA	<i>Emerging Trends in Learning Analytics: Leveraging the Power of Education Data</i>	(Khine, 2019)
2019	EDM	<i>Utilizing Educational Data Mining Techniques for Improved Learning: Emerging Research and Opportunities</i>	(Bhatt, Sajja, & Liyanage, 2019)
2019	LA	<i>Utilizing Learning Analytics to Support Study Success</i>	(Ifenthaler, Mah, & Yau, 2019)
2020	LA	<i>Adoption of Data Analytics in Higher Education Learning and Teaching</i>	(Ifenthaler & Gibson, 2020)
2020	EDM i LA	<i>Big Data on Campus: Data Analytics and Decision Making in Higher Education</i>	(Webber & Zheng, 2020)
2020	LA	<i>Learning Analytics Cookbook: How to Support Learning Processes Through Data Analytics and Visualization</i>	(Jaakonmäki et al., 2020)
2020	LA	<i>Radical Solutions and Learning Analytics: Personalised Learning and Teaching Through Big Data</i>	(Burgos, 2020)
2021	LA	<i>A Beginner's Guide to Learning Analytics</i>	(Srinivasa & Kurni, 2021)
2021	EDM i LA	<i>Advancing the Power of Learning Analytics and Big Data in Education</i>	(Azevedo, Azevedo, Uhoimobhi, & Ossiannilsson, 2021)
2021	LA	<i>Innovative Learning Analytics for Evaluating Instruction: A Big Data Roadmap to Effective Online Learning</i>	(Frick, Myers, Dagli, & Barrett, 2021)
2021	EDM	<i>Modern Approach to Educational Data Mining and Its Applications</i>	(Sweta, 2021)
2021	LA	<i>Visualizations and Dashboards for Learning Analytics</i>	(Sahin & Ifenthaler, 2021)

O analytice učení jako svébytné oblasti zabývající se zkoumáním učení ve virtuálním prostředí se začíná hovořit na přelomu let 2010 a 2011, a to především díky tomu, že v roce 2010 publikoval Siemens na svém blogu zřejmě první definici analytiky učení, která se později stala všeobecně přijímanou v řadě odborných

textů (srov. např. Ferguson, 2012a; Chatti et al., 2012; Siemens, 2013). V roce 2011 pak proběhla první mezinárodní konference zaměřená specificky na analytiku učení (*1st International Conference on Learning Analytics and Knowledge*) a zároveň byla založena společnost pro výzkum v oblasti analytiky učení (*Society for Learning Analytics Research*). Ve stejném roce byla také analytika učení poprvé zmíněna ve výroční zprávě *Horizon Report* (Johnson, Smith, Willis, Levine, & Haywood, 2011), čímž se tato oblast začala dostávat do širšího povědomí odborníků v kontextu e-learningu i vzdělávání obecně.

Analytice učení pak byla věnována výrazná pozornost i v následujících letech. V uvedené výroční zprávě *Horizon Report* se analytika učení nadále objevovala jako jeden z klíčových trendů ve vzdělávání (Johnson, Adams, & Cummins, 2012; Johnson et al., 2013; Johnson, Adams Becker, Estrada, & Freeman, 2014). Od roku 2012 se pak na analytiku učení začaly více zaměřovat i odborné časopisy a postupně vzniklo několik speciálních čísel v prestižních odborných časopisech věnovaných právě tomuto tématu. Jejich základní přehled podává následující tabulka (viz tabulka 2). V roce 2014 pak vychází první číslo nového časopisu zabývajícího se přímo analytikou učení (*Journal of Learning Analytics*) a vychází také první monografie zaměřená specificky na téma analytiky učení (Larsson & White, 2014). V roce 2017 se následně objevila řada dalších publikací zaměřujících se právě na analytiku učení, jak je patrné z tabulky výše (viz tabulka 1).

**Tabulka 2:** Vybraná speciální čísla odborných časopisů zaměřující se na analytiku učení

Rok	Název časopisu	Číslo	Téma speciálního čísla
2012	<i>Journal of Asynchronous Learning Networks</i>	Vol. 16, No. 3	Learning Analytics
2012	<i>Journal of Educational Technology &amp; Society</i>	Vol. 15, No. 3	Learning and Knowledge Analytics
2013	<i>American Behavioral Scientist</i>	Vol. 57, No. 10	Learning Analytics
2013	<i>Journal of the Learning Sciences</i>	Vol. 22, No. 4	Learning Analytics and Computational Techniques for Detecting and Evaluating Patterns in Learning
2015	<i>Journal of Universal Computer Science</i>	Vol. 21, No. 1	Learning Analytics
2015	<i>Computers in Human Behavior</i>	Vol. 47	Learning Analytics, Educational Data Mining and data-driven Educational Decision Making
2016	<i>Journal of Computer Assisted Learning</i>	Vol. 32, No. 3	Learning Analytics in Massively Multi-User Virtual Environments and Courses

Rok	Název časopisu	Číslo	Téma speciálního čísla
2016	<i>Educational Technology Research and Development</i>	Vol. 64, No. 5	Exploring the Relationship of Ethics and Privacy in Learning Analytics and Design: Implications for the Field of Educational Technology
2017	<i>Interactive Learning Environments</i>	Vol. 25, No. 2	Learning Analytics
2017	<i>Zeitschrift für Hochschulentwicklung [Journal for Higher Education Development]</i>	Vol. 12, No. 1	Learning Analytics: Implications for Higher Education
2017	<i>New Directions for Higher Education</i>	Vol. 2017, No. 179	Learning Analytics in Higher Education
2017	<i>Journal of Educational Technology &amp; Society</i>	Vol. 20, No. 2	Learning Analytics in Technology Enhanced Language Learning
2017	<i>ASHE Higher Education Report</i>	Vol. 43, No. 5	Learning Analytics in Higher Education
2017	<i>Journal of Research in Innovative Teaching &amp; Learning</i>	Vol. 10, No. 2	Learning analytics in primary, secondary and higher education
2018	<i>International Journal of Innovation and Learning</i>	Vol. 23, No. 1	User Modelling and Learning Analytics
2018	<i>Journal of Educational Technology &amp; Society</i>	Vol. 21, No. 2	Trends and Research Issues of Learning Analytics and Educational Big Data
2019	<i>Interactive Learning Environments</i>	Vol. 27, No. 5-6	The new potentials for Intelligent Tutoring with learning analytics approaches
2019	<i>User Modeling and User-Adapted Interaction</i>	Vol. 29, No. 4	Learning Analytics and Personalised Support Across Spaces
2020	<i>Australasian Journal of Educational Technology</i>	Vol. 36, No. 6	Learning Analytics: Pathways to Impact

## 2.1 Historické kořeny data miningu ve vzdělávání

Data mining jako takový, někdy označován také jako dobývání znalostí z databází (*knowledge-discovery in databases – KDD*)<sup>4</sup>, se začal výrazněji rozvíjet začátkem 90. let

<sup>4</sup> Nutno podotknout, že ačkoli mnozí autoři dnes chápou pojmy data mining a dobývání znalostí z databází jako vesměs synonymní (Gorunescu, 2011), jiní autoři je rozlišují a chápou data mining pouze jako jeden z kroků v rámci procesu dobývání znalostí z databází. Toto rozlišování obou pojmů vychází především z prací Fayyada (Fayyad, Piatetsky-Shapiro, & Smyth, 1996), podle nějž KDD odkazuje k celkovému procesu získávání užitečných znalostí z dat, kdežto pojem data mining odkazuje pouze ke konkrétnímu kroku v tomto procesu, kdy jsou aplikovány specifické algoritmy pro extrakci určitých vzorců z dat. V této publikaci se přikláním k první skupině autorů a považuji oba pojmy za synonymní.

minulého století a postupem času se začal prosazovat v mnoha oblastech od akademického výzkumu přes soukromou sféru až po různé specifické oblasti využití jako např. lékařství, geografie či internetová bezpečnost.

Data mining lze přitom definovat různými více či méně odlišnými způsoby. Například Gorunescu (2011) uvádí následující možné definice:

- Automatické hledání vzorců (*patterns*) v obrovských databázích s využitím výpočetních technik z oblasti statistiky, strojového učení a rozpoznávání vzorců (*pattern recognition*).
- Netriviální extrakce implicitních, dříve neznámých a potenciálně použitelných informací z dat.
- Věda o získávání užitečných informací z velkých datových souborů nebo databází.
- Proces automatického objevování informací, identifikace vzorců a vztahů „skrytých“ v datech.

Vidíme, že všechny uvedené definice mají jedno společné, totiž že v rámci data miningu jde o práci s velkými datovými soubory, v nichž jsou s využitím specifických výpočetních technik automaticky odhalovány doposud neznámé informace, vzorce či vztahy. Data mining přitom čerpá především ze tří hlavních oblastí, které propojuje a jejichž techniky a metody využívá. V první řadě se jedná o „klasickou“ statistiku, ve které data mining navazuje především na oblast exploratorní analýzy dat<sup>5</sup>. Zde data mining čerpá především základní popisné a explorační statistické a vizualizační techniky. Druhou oblastí, na níž data mining staví, je umělá inteligence (*artificial intelligence*) a strojové učení (*machine learning*). Tyto oblasti propůjčují data miningu především různé techniky zpracování dat založené na pokročilém matematickém aparátu, jakož i sadu algoritmů umožňujících počítači „učit se“ z dat. V neposlední řadě pak data mining navazuje na oblast databázových systémů (*database systems*), kde využívá především technické poznatky o způsobech ukládání a uchovávání dat, struktuře dat a databází, data managementu apod.

Oblast data miningu ve vzdělávání, jakožto oblast zaměřující se na využití data miningu specificky ve vzdělávacím kontextu, pak vznikla především na základě snah o využití data miningu pro analýzu tzv. logů<sup>6</sup> či záznamů o interakci studentů v online vzdělávacích systémech (Baker & Yacef, 2009). Postupně pak byly rozvíjeny stále nové možnosti využití data miningu v oblasti e-learningu a již v roce

---

5 Exploratorní analýzou dat odkazují k rozlišení, které provedl a propagoval Tukey (1962, 1977), jenž se vymezoval vůči klasické statistice a raději používal pojem analýza dat (*data analysis*), který zároveň vnímal jako širší. V rámci analýzy dat pak rozlišoval mezi tzv. konfirmační analýzou dat (*confirmatory data analysis*) a exploratorní analýzou dat (*exploratory data analysis*).

6 Termín „log“ resp. „logy“ má v češtině dva významy. V prvním významu se tímto slovem označuje soubor či tabulka obsahující záznamy o činnosti (např. o činnosti uživatelů určitého systému), v druhém významu se slovem log odkazuje ke každému jednotlivému záznamu o činnosti, který je uložen ve zmiňovaném souboru.

2007 hovoří Castro a spol. (2007) o řadě směrů využívajících data mining v oblasti e-learningu. Zmiňují např. využití data miningu v oblasti hodnocení výkonu studentů, při evaluaci učebních materiálů a online kurzů či data mining zaměřený na detekci atypického učebního chování studentů. Tyto stále se rozšiřující možnosti využití nakonec vedly k ustanovení data miningu ve vzdělávání jakožto svébytné výzkumné oblasti.

## 2.2 Historické kořeny analytiky učení

Analytiku jakožto obecný pojem je poměrně obtížné přesně vymezit. S tím, jak se tento pojem v průběhu posledních patnácti až dvaceti let dostal do širšího povědomí, začal totiž být používán v mnoha různých oblastech a kontextech (Van Barneveld, Arnold, & Campbell, 2012). Pojem analytiky tak najdeme nejen ve spojení s učením (*learning analytics*), ale i ve spojení s řadou dalších oblastí (např. *health analytics*, *news analytics*, *cultural analytics* apod.). Časté je také rozlišování analytiky podle účelu prováděných analýz (např. *predictive analytics*, *prescriptive analytics* či *action analytics*), nebo na základě spojení s konkrétní sadou analytických nástrojů (*google analytics*, *twitter analytics*, *facebook analytics* apod.). Někteří autoři se proto přiklánějí k názoru, že kvůli nadužívání tohoto pojmu je jakýkoli pokus o podání jeho dostatečně podrobné, ale netriviální definice předem odsouzen k nezdaru (Cooper, 2012a).

Pro účely této publikace vycházím z obecně pojaté definice autorek Bienkowski, Feng a Means (2012), které vymezují analytiku v širokém smyslu jako využívání technik z počítačových věd, matematiky a statistiky pro extrakci užitečných informací z velkých objemů dat. Všimněme si přitom tří zásadních aspektů:

1. Analytika čerpá ze znalostí a dovedností jak z oblasti informatiky, tak z oblasti matematiky, a primárně statistiky a analýzy dat.
2. Cílem analytiky je získat informace, které pro nás budou užitečné při dalším plánování či jednání. Jde tedy o informace, jež vedou k nějaké další akci (Cooper, 2012a; Picciano, 2012). Van Barneveld, Arnold a Campbell (2012) dokonce vyzdvihují pouze tento aspekt a definují analytiku jednoduše jako daty řízené rozhodování (*data-driven decision making*).
3. Analytika se zabývá velkými objemy dat, kdy nejsou neobvyklé ani datové sady o desítkách či stovkách milionů záznamů. Zde pak lze spatřovat určitý rozdíl oproti „tradiční“ statistice, která příliš často nepracuje s takto velkými daty.

Do širšího povědomí se pojem analytiky dostal především na základě dvou oblastí: webové analytiky (*web analytics*) a podnikové analytiky (*business analytics* či *business intelligence*). Z nich zároveň čerpaly další specifické oblasti analytiky při vlastním vzniku a formování. Není proto divu, že i u analytiky učení uvádějí mnozí



autoři právě tyto dvě oblasti jako jedny z jejích hlavních historických kořenů (např. Buckingham Shum, 2012; Cooper, 2012b, Ferguson, 2012a, 2012b; Macfadyen & Dawson, 2012).

O webové analytice se začíná hovořit poté, co se rozšířilo komerční využití webových stránek, resp. webových technologií obecně. Jakmile začaly nejrůznější firmy postupně využívat webové stránky ke svému podnikání, ukázalo se jako užitečné (tj. výdělečné) sledovat návštěvnost stránek. Firmy proto začaly věnovat pozornost tomu, jakým způsobem se na jejich webových stránkách uživatelé pohybují a jak je využívají. Na základě těchto znalostí pak firmy mohly nechat své stránky upravit a optimalizovat, což jim následně přinášelo větší zisky. Webová analytika se tak ustanovila jako oblast zabývající se sběrem, měřením, analýzou a vyhodnocováním webových dat za účelem porozumění chování uživatele a optimalizace webových stránek (Kaushik, 2011).

V průběhu 90. let a na přelomu tisíciletí (viz Ferguson, 2012a; Macfadyen & Dawson, 2012) se pak technologie webu začala využívat i pro účely výuky, kdy se mezi vzdělávacími institucemi rozšířilo využívání systémů pro řízení výuky (LMS). Po tomto masovém rozšíření webových výukových systémů se tak i v kontextu vzdělávání začalo uvažovat o možnostech analýzy návštěvnosti online vzdělávacích prostředí s cílem jejich následné optimalizace a zlepšení procesu učení se v nich. Jedním z hlavních pramenů analytiky učení je tak myšlenka aplikace webové analytiky ve specifickém kontextu online vzdělávacích systémů založených na webových technologiích.

Stejně výrazným příspěvkem do formování analytiky učení však byla oblast podnikové analytiky, resp. *business intelligence*. V případě podnikové analytiky jde o využití analytiky především v souvislosti se zvyšováním efektivity firemních procesů a podporou strategického managementu a rozhodování. Tento směr využití analytiky je dokonce starší než webová analytika. Rovněž v kontextu vzdělávání se uvažování o analytice ve vzdělávání ubíralo nejprve tímto směrem. Již v roce 2005 Goldstein a Katz (2005) hovoří o akademické analytice jako o využití *business intelligence* v akademickém sektoru za účelem zlepšení organizačních procesů a zvýšení efektivity v rámci univerzity. Podobně Campbell, DeBlois a Oblingerová (2007) či Fritz (2011) přibližují akademickou analytiku jako oblast zabývající se primárně administrativními a provozními záležitostmi vzdělávacích institucí. Zde může jít například o záležitosti týkající se rozpočtu a plánování, řízení lidských zdrojů, financí a fundraisingu nebo správy institucionálního výzkumu.

Lze tedy říci, že dvě výše zmíněné oblasti využití analytiky ve firemní sféře daly vzniknout dvěma různým směrům využití analytiky ve vzdělávání. Na jedné straně stojí analytika učení inspirována primárně webovou analytikou a zaměřující se na samotné procesy učení a vyučování, na druhé straně se pak nachází akademická analytika, která vychází z podnikové analytiky a souvisí především s problematikou řízení vzdělávacích institucí.

## 2.3 Další oblasti formující analytiku učení a data mining ve vzdělávání

Vedle výše uvedených disciplín či výzkumných oblastí však můžeme identifikovat ještě několik dalších oblastí výzkumu, které se významně podílely na vzniku a formování analytiky učení a data miningu ve vzdělávání. Nejdříve je ale nutno podotknout, že v případě analytiky učení lze za zřejmě nejdůležitější inspirační zdroj považovat právě data mining ve vzdělávání. Jak již bylo zmíněno výše, komunita výzkumníků v rámci data miningu ve vzdělávání se ustanovila o několik let dříve, než vůbec začala být věnována pozornost oblasti označované jako analytika učení. Jakmile se pak začala ustanovovat analytika učení jako svébytná výzkumná oblast, mohla již výrazně čerpat z existující komunity výzkumníků zaměřujících se na data mining ve vzdělávání. Proto je většina níže zmiňovaných oblastí důležitým inspiračním zdrojem nejen pro data mining ve vzdělávání, ale i pro analytiku učení.

Pokud by měl být vybrán jeden hlavní inspirační zdroj pro data mining ve vzdělávání, pak by šlo zřejmě o oblast modelování uživatelů (*user modeling*). Jde o oblast výzkumu v rámci multidisciplinárního oboru studujícího interakci mezi člověkem a počítačem<sup>7</sup>. Cílem výzkumníků v této oblasti je porozumět tomu, jak různí uživatelé interagují s určitou aplikací či uživatelským rozhraním. Následně je možné vytvořit model, na jehož základě může daná aplikace či rozhraní reagovat odlišnými způsoby na různé formy interakce uživatele s aplikací či rozhraním. A to například tím, že se vzhled aplikace přizpůsobí podle potřeb uživatele nebo je uživateli doporučen pro něj relevantní obsah apod. V kontextu data miningu ve vzdělávání se pak často hovoří o modelování studentů (viz např. Romero, Ventura, Pechenizkiy, & Baker, 2010; Peña-Ayala, 2014; Papamitsiou & Economides, 2014), kam může spadat např. modelování chování studentů, modelování jejich emočních či afektivních stavů, modelování motivace či nálady studentů apod. Dále je třeba doplnit, že řada autorů (např. Ferguson, 2012a, 2012b; Chatti et al., 2012; Siemens, 2013) zmiňuje modelování uživatelů zároveň jako jeden z důležitých inspiračních zdrojů i v souvislosti s formováním analytiky učení.

Zásadní důležitost široké oblasti uživatelského modelování pro data mining ve vzdělávání i analytiku učení lze spatřovat především v tom, že určitá forma modelování uživatelů se stala jejich inherentní součástí. Řada projektů a výzkumů v oblasti analytiky a data miningu ve vzdělávání tak zahrnuje alespoň nějaký aspekt modelování uživatelů. Zároveň je vliv modelování uživatelů patrný také v tom, že vedle vlastního modelování uživatelů existuje hned několik specifitější zaměřených výzkumných oblastí, které jsou na modelování uživatelů postaveny a které lze rovněž považovat za inspirační zdroje analytiky a data miningu ve vzdělávání.

---

7 V angličtině se pro tento obor vžilo označení *Human-Computer Interaction* (HCI), které se do češtiny překládá jen velmi zřídka.

Předně lze zmínit oblast doporučovací systémů (*recommender systems*), kterou v kontextu analytiky i data miningu ve vzdělávání zmiňuje řada autorů (např. Romero & Ventura, 2013; Siemens, 2013; Peña-Ayala, 2014a, 2014b). Doporučovací systémy jsou samozřejmě široce používány i mimo vzdělávací prostředí<sup>8</sup>. V kontextu vzdělávání je pak cílem těchto systémů doporučit studentům takový výukový obsah (nebo kurz, úkol, odkaz apod.), který je pro ně v dané době nejvhodnější. Právě zde vstupuje do hry již zmiňovaná problematika modelování uživatelů. Chceme-li totiž doporučit studentům vhodný obsah například na základě jejich znalostí, je nutné mít nejprve nějaký model studentů s ohledem na jejich znalosti. Teprve na základě vytvořeného modelu jim pak lze nabízet odpovídající obsah. Základní fungování doporučovací systémů lze tudíž popsat tak, že nejprve shromažďují potřebné údaje o chování, znalostech či preferencích studentů, a s využitím modelování pak z těchto dat vyvozují závěry ohledně obsahu, který by byl pro konkrétního studenta vhodný či který by jej mohl zajímat apod. (viz Chatti et al., 2012).

K dalším oblastem, jež jsou do značné míry založeny na modelování uživatelů, patří oblast inteligentních tuteurských systémů (*intelligent tutoring systems*) a oblast adaptivních hypermediálních systémů (*adaptive hypermedia systems*). V prvním případě jde o využívání (především) kognitivního modelování ve spojení s umělou inteligencí, přičemž cílem je tvorba takových vzdělávacích systémů, které provází studenta v průběhu učení a jsou schopny mu poskytovat okamžitou pomoc a personalizovanou zpětnou vazbu. V druhém případě jde o systémy založené na hypertextovém a hypermediálním obsahu, jež jsou schopny s využitím modelování uživatelů tento obsah adaptovat a přizpůsobit jej potřebám jednotlivých studentů (Brusilovsky, 2001). Rozvoj obou těchto oblastí byl zásadním impulzem především v souvislosti se vznikem data miningu ve vzdělávání (ale částečně také se vznikem analytiky učení, viz např. Chatti et al., 2012; Siemens, 2013). Tyto oblasti se rozvíjejí i v současnosti a nadále dochází k výrazným styčným plochám mezi nimi a data miningem ve vzdělávání.

Někteří autoři z oblasti data miningu ve vzdělávání (např. Baker & Yacef, 2009; Bousbia & Belamri, 2014; Romero & Ventura, 2010, 2013) rovněž poukazují na psychometriku (*psychometrics*) jako na jeden z významných historických kořenů data miningu ve vzdělávání. Jak uvádějí Urbánek, Denglerová a Širůček (2011, s. 13), „psychometrika je nejčastěji chápána jako jakási pomocná psychologická disciplína zaměřená na tvorbu, adaptaci, úpravy a hodnocení psychodiagnostických metod“. Zároveň tito autoři doplňují, že je nutné psychometriku chápat spíše jako oblast zabývající se obecně teoretickými otázkami měření v psychologii, přičemž samotná praktická aplikace určitých matematických či statistických postupů je v tomto

---

8 Čtenář se jistě setkal např. s doporučovacím systémem (resp. jeho výstupy) na platformě YouTube, kde je jeho cílem doporučit uživatelům relevantní videa ke zhlédnutí.

smyslu až druhotná, jelikož vždy navazuje na obecné teoretické principy. Vliv psychometrie na formování data miningu ve vzdělávání lze přitom vidět především v oblasti metodologie, neboť řada autorů v komunitě data miningu ve vzdělávání využívá postupů z oblasti psychometrického modelování. Jako příklad lze zmínit teorii odpovědi na položku (*item response theory – IRT*) či využití bayesovských sítí (Baker & Yacef, 2009). Romero a Ventura (2010) dokonce zmiňují, že přístupy využívané v oblasti data miningu ve vzdělávání a stavějící na psychometrickém modelování, jsou pro obecný data mining poměrně neobvyklé a nejsou univerzálně chápány jakou součástí data miningu. Je tedy patrné, že právě tyto metodologické inspirace v psychometrii jsou specifické právě pro data mining ve vzdělávání. Využívání do určité míry stejných metod je pak zřejmě odrazem toho, že se do komunity výzkumníků v oblasti data miningu ve vzdělávání zapojila řada výzkumníků původně se pohybujících v oblasti psychometrie (Baker & Yacef, 2009). Zároveň jsou zde patrné přesahy v kontextu měření znalostí a tvorby didaktických testů, čemuž je v psychometrii dlouhodobě věnována značná pozornost. Není tak vůbec od věci, že Baker, Martin a Rossi (2017) chápou oblast data miningu ve vzdělávání jako oblast, která do značné míry navazuje na relativně dlouhou historii problematiky měření v psychologii.

Co se týče dalších metodologických inspirací, upozorňují Chatti et al. (2012) či Chernobilsky et al. (2016) na výrazné styčné plochy mezi analytikou učení a akčním výzkumem (*action research*). Základní charakteristiky akčního výzkumu a analytiky učení (ale částečně i data miningu ve vzdělávání) jsou totiž dle těchto autorů velmi podobné. Akční výzkum v pedagogice se často zaměřuje na konkrétní školní situaci a není neobvyklé, když se na něm podílí samotní učitelé. Jeho hlavním cílem je zdokonalování pedagogické praxe a zvyšování kvality práce učitelů, což má vést k celkovému zkvalitňování poskytovaného vzdělávání. Esenciální charakteristikou je zásah do dosavadní vzdělávací praxe. V průběhu akčního výzkumu proto dochází k zavádění určitých změn. Prováděné intervence mají cyklický charakter, kdy je každá z intervencí nejprve vyhodnocena a teprve poté je stanovován další postup výzkumu. Průběh akčního výzkumu tak bývá často konceptualizován v podobě iterativního cyklu (Nezvalová, 2003). Stejně tak postup v analytice učení bývá pojímán jako opakující se cyklus zahrnující průběžné intervence a zavádění změn (viz kapitolu 4.1). Podobně je cílem analytiky učení zlepšení a zkvalitnění edukační praxe.

Přesto můžeme mezi akčním výzkumem a analytikou učení nalézt i několik zásadních rozdílů. Zatímco na akčním výzkumu často participují především učitelé a studenti (resp. žáci), výzkumné projekty založené na analytice učení často zahrnují více zúčastněných stran (např. vedení instituce, technické oddělení instituce apod.). Chatti et al. (2012) zmiňují rozdíl také v tom, čím vlastně samotný výzkumný projekt začíná. U akčního výzkumu jde obvykle o problémy či otázky, které vyplývají z vlastní pedagogické praxe. Výzkumné projekty v rámci analytiky

učení se naopak formují často až na základě pozorování a analyzování již získaných dat. V neposlední řadě se analytika učení a akční výzkum odlišují v samotných výzkumných metodách a postupech, které používají. Zatímco v akčním výzkumu je běžnější použití kvalitativních metod, analytika učení stojí téměř výhradně na metodách kvantitativních.

V neposlední řadě je třeba spolu se Siemensem (2013) zmínit jako důležitý zdroj mnoha inspirací pro analytika učení (ale i data mining ve vzdělávání) celou oblast e-learningu. A to ve smyslu, jež v českém kontextu dlouhodobě prosazuje především Zounek a jeho žáci, kteří chápou e-learning jako širokou oblast jak teoretického studia a empirických výzkumů, tak i reálných vzdělávacích procesů spočívajících na práci s informačními a komunikačními technologiemi (viz Zounek, 2009 či nověji Zounek et al., 2016). Právě obrovské rozšíření a reálné využívání ICT ve výuce totiž umožnilo vznik oblastí, jako jsou data mining ve vzdělávání a analytika učení. Zároveň široké spektrum výzkumů realizovaných v oblasti e-learningu výrazným způsobem přispělo ke znalostní i metodologické bázi obou těchto oblastí, a ovlivnilo tak jejich další formování a směřování.