

Juhaňák, Libor

Vymezení analytiky a data miningu ve vzdělávání a jejich rozdíly

In: Juhaňák, Libor. *Analytika učení a data mining ve vzdělávání v kontextu systémů pro řízení výuky*. Vydání první Brno: Masarykova univerzita, 2023, pp. 25-34

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

Stable URL (handle): <https://hdl.handle.net/11222.digilib/digilib.77691>

Access Date: 22. 02. 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.

3 VYMEZENÍ ANALYTIKY A DATA MININGU VE VZDĚLÁVÁNÍ A JEJICH ROZDÍLY

Po nastínění historického kontextu, který byl nabídnut v předchozí kapitole, se tato kapitola zaměřuje na přesnější vymezení oblastí označovaných jako data mining ve vzdělávání a analytika učení. Nejprve je blíže vymezen data mining ve vzdělávání a následně je rozebrána definice analytiky učení. Nakonec je věnována podrobnější pozornost srovnání obou oblastí, a to především s ohledem na základní odlišnosti mezi těmito výzkumnými směry.

3.1 Vymezení data miningu ve vzdělávání

Ačkoli v rámci data miningu ve vzdělávání nenajdeme jednu všeobecně přijímanou definici této výzkumné oblasti, za jednu ze stěžejních definic lze považovat tu, kterou nabízejí Baker a Yacefová (2009). Ti ve své přehledové studii vymezují oblast data miningu ve vzdělávání jako „vznikající disciplínu zabývající se rozvojem metod pro zkoumání jedinečných typů dat, která pocházejí ze vzdělávacích prostředí, a použití těchto metod pro lepší porozumění studentům a prostředím, ve kterých se učí“⁹. Toto vymezení data miningu ve vzdělávání je zároveň přijímáno i Mezinárodní společností pro data mining ve vzdělávání a souvisejícím časopisem, který tato společnost vydává¹⁰.

9 Originální znění: „Educational Data Mining is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in.“ (Baker & Yacef, 2009).

10 Jak již bylo zmíněno výše, jde o společnost s názvem *International Educational Data Mining Society* a časopis nazvaný *Journal of Educational Data Mining*.

3 Vymezení analytiky a data miningu ve vzdělávání a jejich rozdíly

Výše uvedená definice je však poměrně široká a zároveň poněkud překvapivě vůbec nezmiňuje obecnou oblast data miningu. Naopak Romero a Ventura (2013) se k data miningu jako takovému výslovně hlásí, když říkají, že oblast data miningu ve vzdělávání může být jednoduše definována jako „aplikace data miningových technik na specifické typy dat, pocházející ze vzdělávacích prostředí“¹¹. Přitom podobně jako Baker a Yacefová (2009) upozorňují, že cílem data miningu ve vzdělávání je hledání odpovědí na důležité pedagogické otázky. Bakhshinategh, Zaiane, ELAtia a Ipperciel (2018) rovněž odkazují k obecnému data miningu a definují data mining ve vzdělávání jednoduše jako oblast využívající techniky data miningu ve vzdělávacích prostředích. Romero a Ventura (2013) následně vymezení oblasti data miningu ve vzdělávání zpřesňují a tvrdí, že se „zabývá vývojem, výzkumem a aplikací počítačových metod pro detekci vzorců ve velkých sbírkách vzdělávacích dat, která by jinak byla analyzovatelná jen velmi těžko či vůbec“¹². Podobně Baker a Inventado (2014) upřesňují, že v případě data miningu ve vzdělávání jde především o velké objemy dat (*large-scale educational data*) a o aplikaci automatických metod (tj. metod z oblasti data miningu) při jejich analýze.

Calders a Pechenizkiy (2011) kladou při vymezení data miningu ve vzdělávání důraz na poněkud odlišné prvky než výše zmiňovaní autoři. Předně považují data mining ve vzdělávání za multidisciplinární výzkumnou oblast, do které vstupují výzkumníci a specialisté z disciplín, jako jsou informatika, pedagogika, psychologie, psychometrika a statistika. Dále zdůrazňují, že data mining ve vzdělávání je zároveň vědou o učení¹³ i aplikační oblastí obecného data miningu. Vždy je však podle uvedených autorů primárním cílem porozumění tomu, jak probíhá učení. V neposlední řadě pak Calders a Pechenizkiy vidí data mining ve vzdělávání jako oblast, která umožňuje daty řízené rozhodování, s jehož využitím pak usiluje o zlepšení současné vzdělávací praxe. Tím data mining ve vzdělávání do značné míry přibližují akčnímu výzkumu (podobně jako je tomu u analytiky učení).

Kdybychom chtěli podat určité shrnutí různých definic data miningu ve vzdělávání a nabídnout obecné vymezení této oblasti, pak lze za hlavní charakteristické prvky data miningu ve vzdělávání považovat následující:

11 Originální znění: „In fact, EDM, can be defined as the application of data mining (DM) techniques to this specific type of dataset that come from educational environments to address important educationa questions.“ (Romero & Ventura, 2013).

12 Originální znění: „Educational data mining (EDM) is concerned with developing, researching, and applying computerized methods to detect patterns in large collections of educational data that would otherwise be hard or impossible to analyze (...).“ (Romero & Ventura, 2013).

13 Zde zmiňovaní autoři využívají termín „learning science“, čímž podle všeho odkazují k oblasti označované v zahraniční literatuře jako *learning sciences*. Vzhledem k zaměření této publikace není nutné věnovat pozornost tomu, nakolik jde v případě věd(y) o učení o zvláštní výzkumné pole a nakolik jde o jiné označení či součást běžného pedagogického výzkumu. Pro podrobnější informace lze proto pouze odkázat na stránky Mezinárodní společnosti pro vědy o učení (*International Society of the Learning Sciences*) na adrese www.isls.org.

- Jde o aplikaci metod a technik (obecného) data miningu v kontextu vzdělávání.
- Pozornost je věnována především „velkým datům“ (*big data*), která jsou obtížně analyzovatelná běžnými statistickými metodami.
- Cílem je s využitím zmiňovaných metod a dat získávat odpovědi na pedagogické otázky týkající se učení a prostředí, v nichž učení probíhá.
- Jde o nově se formující oblast čerpající teoretické a metodologické přístupy i z dalších relevantních oblastí mimo obecný data mining.

3.2 Vymezení analytiky učení

Analytiku učení lze vymežit různými způsoby. Asi nejčastěji přijímanou definicí je definice Siemense, který chápe analytiku učení jako disciplínu zabývající se „měřením, sběrem, analýzou a reportováním dat o studentech a jejich kontextech za účelem pochopení a optimalizování učení a prostředí, ve kterých učení probíhá“ (Long & Siemens, 2011, s. 34)¹⁴. Jak je patrné, tato definice zřetelně navazuje na tradici webové analytiky, když se zaměřuje přímo na vzdělávací prostředí a jeho optimalizaci. Na druhou stranu však odhlíží od technických specifik a zdůrazňuje výhradně pedagogickou stránku. Nespojuje rovněž analytiku učení s dříve ustanovenými konkrétními oblastmi, nýbrž ji vymezuje jako svébytnou disciplínu.

Zároveň je třeba zdůraznit, že Siemensova definice neomezuje analytiku učení pouze na online prostředí, jako to dělají jiní autoři. Například Chatti et al. (2012) či Fergusonová (2012a) chápou analytiku učení jako do značné míry dílčí výzkumnou oblast v rámci e-learningu, která se týká výhradně učení v online prostředí. Takové pojetí analytiky učení je však v dnešní době možné pokládat za příliš úzké. Ačkoli se zřejmě většina výzkumníků i dnes věnuje analytice učení v online vzdělávacích systémech, existují již příklady využití analytiky učení i v offline vzdělávacích prostředích (např. ve školní třídě). Autoři jako D’Mello (2017) či Ochoa (2017) zmiňují příklady výzkumů toho, kam se studenti v průběhu výuky dívají, analýzy postojů, gest a pohybů či analýzy výrazů tváře studentů v průběhu učení či při plnění vzdělávací aktivity. V tomto kontextu se pak hovoří o oblastech jako *Emotional Learning Analytics*, *Classroom Learning Analytics* či *Multimodal Learning Analytics*. Z tohoto pohledu se tedy jeví Siemensova širěji pojatá definice analytiky učení jako vhodnější.

Na druhou stranu však výše uvedená Siemensova definice postihuje pouze jednu z úrovní, o kterých je možné v souvislosti s analytikou ve vzdělávání uvažovat.

14 Originální znění: „learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs.“ (Long & Siemens, 2011, s. 34).

Podle Buckingham Shuma (2012) lze přitom rozlišovat tři základní úrovně analytiky ve vzdělávání:

1. Mikroúroveň se zaměřuje na individuální studenty či skupiny studentů¹⁵. Na této úrovni jde o sběr a analýzu dat, která za sebou nechávají jednotliví uživatelé při používání různých systémů či služeb. Nejčastěji jde buď o záznamy z různých vzdělávacích informačních systémů (*Educational Information System, Student Information System* apod.), zahrnujících například absolvované předměty či získané známky, anebo o záznamy aktivit studentů v online výukovém prostředí (např. systémy typu LMS). Taková data přitom umožňují zjistit, jaké studijní materiály si student zobrazil, jak dlouho mu trvalo nějaké cvičení apod. Může však jít i o další typy dat týkajících se studentů, např. výpůjčky v knihovně, sociální síť kontaktů, záznamy o pohybu studenta získané pomocí geolokace aj. Výsledky analýz na mikroúrovni jsou přínosné především pro učitele. Mohou však být užitečné i pro samotné studenty, kterým dovedou poskytovat vhled do vlastních (třeba neuvědomovaných) učebních návyků a podporovat u nich reflexi vlastního učení.
2. Mezoúroveň je úroveň vzdělávacích institucí, na níž je hlavním cílem analytiky učení optimalizace administrativních a organizačních procesů v rámci vzdělávací instituce. Data pocházejí primárně z nejrůznějších ekonomických, správních či manažerských systémů, často však bývají doplňována o data z dalších specifických systémů, které daná instituce využívá. V případě univerzit jsou to zvláště systémy schraňující data o výzkumné a publikační činnosti apod. Výsledky analytiky na mezoúrovni mohou být přínosné nejen pro administrativu, management a vedení vzdělávací instituce, ale také pro investory či donátory a další zainteresované strany.
3. Makroúroveň pak odpovídá regionální, státní, národní či mezinárodní úrovni. Zde tedy jde spíše o meziinstitucionální či nadinstitucionální úroveň, kde analytika může hrát důležitou roli jednak pro jednotlivé vzdělávací instituce, jednak pro nejrůznější regionální a státní školské orgány, národní vlády či mezinárodní organizace.

Na základě výše uvedeného členění analytiky ve vzdělávání je patrné, že Siemensem navrhovaná definice odpovídá mikroúrovni. Je proto nutné rozlišovat užší a širší význam pojmu „analytika učení“. V užším významu se pohybujeme pouze na mikroúrovni, v širším významu pod tento pojem spadají všechny uvedené úrovně. Z důvodu přesnosti by proto bylo vhodnější oba významy explicitně rozlišovat (např. analytika učení vs. analytika vzdělávání). V dosavadních zdrojích však toto rozlišení není dodržováno a využívá se pouze označení analytika učení (*learning analytics*).

15 Hovořím zde sice primárně o studentech, uvedené se však může týkat i žáků.

Zároveň vidíme, že mezoúroveň do značné míry odpovídá tomu, co bylo výše identifikováno jako jedna ze souvisejících oblastí, totiž akademická analytika. Stejně tak Long a Siemens (2011) rozlišují mezi akademickou analytikou a analytikou učení právě na základě toho, že akademická analytika se zaměřuje na úroveň institucí a výše (tj. mezoúroveň plus makroúroveň), zatímco analytika učení zahrnuje úroveň jednotlivých kurzů (tj. mikroúroveň). V takovém případě můžeme analytiku vzdělávání (tj. analytiku učení v širokém smyslu) chápat jako spojení akademické analytiky a analytiky učení v úzkém smyslu.

Zároveň je však třeba upozornit, že výše uvedené pojmy a jejich rozlišení nejsou zatím široce ukotveny a je možné, že se budou postupem času dále měnit, zpřesňovat či rozšiřovat. Například označení „akademická analytika“ jasně odkazuje k místu svého vzniku, totiž k akademickému prostředí univerzit. Lze ovšem očekávat rozšíření analytiky i mimo univerzitní prostředí, proto je možné, že se v budoucnu rozšíří obecnější pojem, který by nezahrnoval jen akademickou sféru (např. institucionální analytika). Zároveň se již dnes hovoří o analytice učení v kontextu profesního vzdělávání a informálního učení (viz např. Littlejohn, 2017), přičemž se využívá označení *Professional Learning Analytics*. V časopise *Journal of Research in Innovative Teaching & Learning* pak v roce 2017 vyšlo speciální číslo zaměřující se mimo jiné na analytiku učení v kontextu primárního a sekundárního vzdělávání (název čísla zněl *Learning analytics in primary, secondary and higher education*, viz tabulka 1). Lze proto očekávat, že s rozšiřováním analytiky učení do nových oblastí a kontextů se budou objevovat i další specifická označení konkrétních směrů výzkumu a aplikace analytiky ve vzdělávání.

3.3 Odlišnosti mezi analytikou učení a data miningem ve vzdělávání

V předchozích dvou kapitolách bylo podáno základní vymezení data miningu ve vzdělávání a analytiky učení. Přestože byly obě tyto oblasti popsány a vymezeny zvlášť a se zaměřením na různé aspekty, byly čtenáři jistě patrné značné podobnosti a výrazné styčné plochy mezi oběma výzkumnými oblastmi. V obou případech jde o relativně mladá výzkumná pole, která se zaměřují na využití dat z různých typů vzdělávacích systémů, přičemž hlavní pozornost je věnována online výukovým systémům. Pro analýzu těchto dat přitom obě oblasti využívají do značné míry obdobné metody, techniky a analytické postupy. Data mining ve vzdělávání i analytika učení se tak zaměřují na využití obdobných typů dat pocházejících z obdobných typů vzdělávacích systémů za využití obdobných metod analýzy. V neposlední řadě je i základní cíl obou oblastí do značné míry podobný. Obě disciplíny totiž usilují jednak o využití výše zmiňovaných dat a analytických postupů pro pochopení a porozumění procesu učení v různých vzdělávacích prostředích a systémech, jednak o následné optimalizování daných výukových

3 Vymezení analytiky a data miningu ve vzdělávání a jejich rozdíly

systémů tak, aby se v nich studenti mohli efektivněji učit a učitelé v nich mohli lépe vyučovat.

Podobnosti mezi analytikou učení a data miningem ve vzdělávání se tak zdají být na první pohled poměrně zřejmé. Co však příliš zřejmé být nemusí, jsou odlišnosti mezi těmito výzkumnými oblastmi, tj. to, zda se tyto oblasti nějakým zásadním způsobem liší, a pokud ano, na jakém základě je možné odlišit jednu od druhé. Následující podkapitoly se proto zaměřují na hlavní rozdíly, které lze mezi analytikou učení a data miningem ve vzdělávání nalézt.

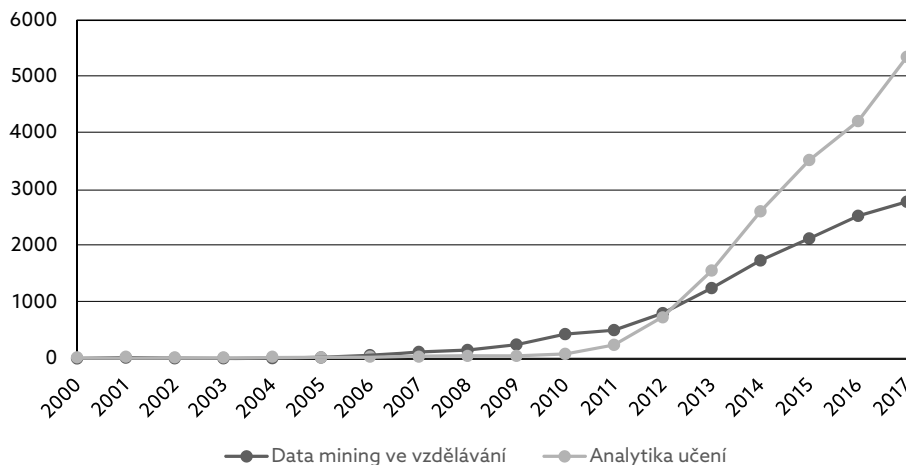
3.3.1 Rozdílný vznik a vývoj v čase

Vzniku a historickým kořenům data miningu ve vzdělávání a analytiky učení již byla věnována podrobná pozornost v kapitole 2. Lze proto jen shrnout, že obě disciplíny vychází z poněkud odlišného kontextu a na jejich formování se podílely jiné oblasti. Zatímco data mining ve vzdělávání vychází primárně z obecného data miningu a databázových systémů, přičemž má blízko rovněž k oblasti umělé inteligence a oblasti strojového učení, analytika učení čerpá především z webové analytiky a podnikové analytiky, resp. z *business intelligence*. Siemens a Baker (2012) doplňují, že data mining ve vzdělávání má blízko především k vývoji vzdělávacího softwaru a modelování uživatelů (studentů), kdežto analytika učení spojuje spíše s oblastí sémantického webu a problematikou tzv. inteligentního kurikula (*intelligent curriculum*) a systemických intervencí (*systemic interventions*).

Do značné míry odlišná je i doba vzniku, a především pak následující vývoj obou oblastí v čase. Jak již bylo zmíněno, vznik data miningu ve vzdělávání jakožto svébytné výzkumné oblasti bývá spojován s přelomem let 2008 a 2009. Diskuze a určité výzkumné snahy lze však v souvislosti s touto oblastí nalézt přinejmenším již v roce 2005, přičemž dílčí pokusy o využití data miningu v kontextu vzdělávání se objevují i hlouběji v minulosti (Romero & Ventura, 2007). Naopak pojem analytiky učení, odkazující k nově se vyčleňující výzkumné oblasti, se objevil poměrně náhle na přelomu let 2010 a 2011. Analytika učení se tak začala rozvíjet o několik let později, a i v současnosti je stále poměrně mladým výzkumným polem.

I přes svůj pozdější vznik však analytika učení zažívá výrazně mohutnější rozvoj než data mining ve vzdělávání. To ilustrují následující dva grafy zaměřující se na četnost výskytu jednotlivých disciplín (tj. frází „learning analytics“, resp. „educational data mining“) v odborných publikacích. Konkrétně byl využit webový vyhledávač odborných textů Google Scholar, v jehož rámci byly pro roky 2000 až 2017 vyhledány počty publikací, v nichž se kdekoli v textu dokumentu (obrázek 1) a následně pouze v názvu dokumentu (obrázek 2) vyskytují klíčová slova „learning analytics“, resp. „educational data mining“.

Počty nalezených výsledků zmiňujících analytiku učení a data mining ve vzdělávání kdekoli v dokumentu



Obrázek 1: Počty nalezených výsledků v rámci Google Scholar zmiňujících analytiku učení a data mining ve vzdělávání dle jednotlivých let¹⁶

Z obou grafů je patrné, že o oblasti data miningu ve vzdělávání se v odborných publikacích začalo hovořit dříve. Jakmile se však kolem let 2010/2011 objevuje analytika učení, tak okamžitě začíná přitahovat pozornost výzkumníků, a po roce 2012 už z hlediska počtu zmínek v relevantní odborné literatuře převyšuje oblast data miningu ve vzdělávání. Přestože jde jen o ilustrační čísla, trend rozdílného vývoje obou oblastí je patrný. Budeme-li vycházet z výskytu označení jednotlivých disciplín v názvech odborných publikací (viz obrázek 2), pak lze říci, že k roku 2017 je „popularita“ analytiky učení mezi výzkumníky několikanásobně vyšší ve srovnání s oblastí data miningu ve vzdělávání. K obdobným záměrům docházejí i Calvet Liñán a Juan Pérez (2015), kteří ovšem tento vývoj mapují pouze do roku 2014.

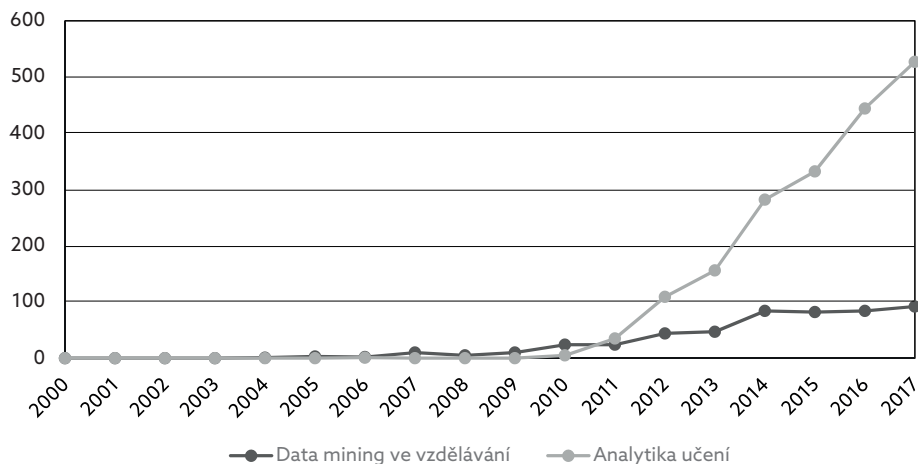
Mimo dále zmiňované odlišnosti může k vyšší popularitě analytiky učení přispívat právě to, jakým způsobem vznikala, resp. které další oblasti se podílely na jejím formování. Zatímco oblast data miningu ve vzdělávání vychází především z obecného data miningu, jež lze považovat do značné míry za technický obor¹⁷, na formování analytiky učení se podílelo širší spektrum dalších i netechnicky orientovaných oblastí. Zřejmě i díky tomu je dnes analytika učení přístupnější výzkumníkům z různých vědních disciplín, a tudíž je ve větším množství zmiňována

¹⁶ Výskyt hledaného termínu „learning analytics“, resp. „educational data mining“ není omezen (tj. může se objevit kdekoli v dokumentu).

¹⁷ Data mining bývá obvykle chápán jako dílčí disciplína či výzkumné pole v rámci informatiky (*computer science*).

3 Vymezení analytiky a data miningu ve vzdělávání a jejich rozdíly

Počty nalezených výsledků zmiňujících analytiku učení a data mining ve vzdělávání v názvu dokumentu



Obrázek 2: Počty nalezených výsledků v rámci Google Scholar zmiňujících analytiku učení a data mining ve vzdělávání dle jednotlivých let¹⁸

v odborných publikacích. S tím souvisí také fakt, že v rámci analytiky učení rovněž působí vyšší počet netechnicky zaměřených výzkumníků, například výzkumníci pocházející primárně z oblasti sociálních věd jako psychologie, sociologie či pedagogika (srov. Bienkowski, Feng, & Means, 2012).¹⁸

3.3.2 Rozdílné využívání analytických metod a technik

Ačkoli se používané analytické metody a techniky v rámci data miningu ve vzdělávání a analytiky učení do značné míry překrývají, můžeme zde nalézt určité odlišnosti v důrazu, který je přikládán jednotlivým metodám a technikám (srov. Baker & Inventado, 2014; Romero & Ventura, 2013; Siemens & Baker, 2012). Pokud bychom vyšli ze základního přehledu používaných metod, jak je prezentován v jedné z následujících kapitol (viz kapitolu 5), pak lze říci, že v oblasti data miningu ve vzdělávání převažuje používání prediktivních metod a technik, zvláště pak technik klasifikace. Dále je typické využívání široké škály různých technik shlukování a také metod dolování vztahů a objevování pomocí modelů. Naopak v analytice učení jsou dominantní metody jako analýza sociálních sítí či různé postupy založené

¹⁸ Výskyt hledaného termínu „learning analytics“, resp. „educational data mining“ je omezen pouze na název dokumentu.

na zpracování přirozeného jazyka (např. analýza sentimentu, analýza diskurzu, analýza obsahu apod.). V rámci analytiky učení jsou rovněž častěji využívány metody a techniky tradiční statistiky.

Data mining ve vzdělávání je více zaměřen na rozvoj technik pro získávání užitečných informací z dat a na vývoj nových algoritmů či modelů, zatímco v kontextu analytiky učení je častěji věnována pozornost aplikaci již existujících postupů v konkrétních vzdělávacích kontextech. S tím souvisí také to, že v analytice učení je kladen větší důraz na analýzu a popis samotných dat, zatímco v data miningu ve vzdělávání se výzkumníci častěji věnují popisu a srovnávání různých technik za účelem identifikace těch, které si v daném kontextu či při řešení daného problému vedou nejlépe.

Výše nastíněný důraz na určité skupiny metod a technik a důraz na popis samotných dat versus rozvoj metod analýzy je mimo jiné reflektován v tom, jakou roli ve zmíněných oblastech hraje lidský úsudek, nebo naopak automatizace. Analytika učení je více zaměřená na tzv. daty řízené rozhodování (*data-driven decision-making*), kde je cílem využít možnosti analýzy dat pro získání informací potenciálně důležitých pro účely rozhodování. Při takovém rozhodování se uplatňuje především lidský úsudek, zatímco různé automatizované techniky se mohou uplatňovat spíše v rámci získávání informací z dat. Čili automatizované techniky analýzy dat jsou pouze prostředkem k tomu, aby bylo možné provádět lepší rozhodování na základě lidského úsudku. Naopak v rámci data miningu ve vzdělávání je automatizace často samotným cílem, kdežto lidský úsudek je prostředkem, jak se k tomuto cíli dopracovat. Tzn. že cílem je dosáhnout automatizace určitého úkolu (např. doporučení vhodného obsahu ke studiu, volba otázky odpovídající náročnosti v rámci adaptivního testovacího systému apod.). Lidský úsudek je pak při dosahování tohoto cíle zapojen například při volbě přístupu k řešení daného problému, při volbě analytických metod atd.

Obecně lze říci, že v rámci data miningu ve vzdělávání je kladen větší důraz na rozvoj a používání metod, které mohou být využity pro účely vývoje prediktivních a adaptivních systémů. Naopak analytika učení akcentuje ty metody, jež jsou přínosné z hlediska poskytnutí vhledu do dat. Analytika učení tak cílí především k tomu, aby za využití různých analytických metod a technik bylo možné poskytovat potenciálně důležité informace jednotlivým zúčastněným stranám (studentům, učitelům, vedení vzdělávací organizace, administrátorům daného online vzdělávacího prostředí apod.).

3.3.3 Rozdílné zaměření pozornosti

S rozdílným vznikem a různými historickými kořeny stejně jako s odlišným důrazem na využívané analytické postupy souvisí i rozdílný způsob, jakým výzkumníci v data miningu ve vzdělávání a v analytice učení zaměřují svou pozornost při

řešení výzkumných problémů a otázek. Siemens a Baker (2012) v této souvislosti hovoří o rozdílném převažujícím přístupu či paradigmatu, které je v těchto oblastech uplatňováno a které v jedné či druhé oblasti převažuje.

Dle Siemense a Bakera (2012) převažuje v oblasti data miningu ve vzdělávání spíše redukcionistický přístup, při němž výzkumníci postupují tak, že redukuje či zjednodušují zkoumané komplexní jevy na dílčí, a co možná nejjednodušší části, které daný jev tvoří. Následně pak věnují pozornost těmto jednotlivým částem a analyzují vztahy mezi nimi. Jako typický příklad takového redukcionistického přístupu uvádějí zmiňovaní autoři níže popisovanou metodu objevování pomocí modelů (viz kapitolu 5.7). Oproti tomu v analytice učení je podle Siemense a Bakera (2012) běžnější holistický přístup, kdy výzkumníci usilují o porozumění zkoumanému jevu jakožto celku. Holistický přístup obecně staví na tom, že systém jako celek nelze zcela vysvětlit zaměřením se pouze na jeho jednotlivé části. Proto i v rámci analytiky učení je mnohem více zdůrazňován širší kontext zkoumaného problému či jevu. Zároveň jsou zohledňovány různé úrovně, kterých se zkoumaný problém dotýká, příp. je výzkumný problém řešen s ohledem na různé zúčastněné strany.

I přes všechny výše zmiňované odlišnosti je však třeba mít na paměti, že se stále jedná o do značné míry překrývající se výzkumné oblasti. Jak upozorňují Siemens a Baker (2012) či Baker a Inventado (2014), výše uváděné odlišnosti odpovídají spíše širším trendům či přístupům v analytice učení a data miningu ve vzdělávání. Nelze je tedy chápat jako definiční charakteristiky jednotlivých oblastí. Řada výzkumníků z oblasti analytiky učení může provádět výzkum, jenž by na základě výše uvedených odlišností mohl lépe zapadat do oblasti data miningu ve vzdělávání, a naopak. Nastíněné odlišnosti je tak nutno chápat spíše jako obecné zaměření jednotlivých oblastí, které však nemusí být platné na úrovni konkrétních výzkumníků či konkrétních realizovaných výzkumů.