

E-LOGOS

ELECTRONIC JOURNAL FOR PHILOSOPHY

ISSN 1211-0442

19/2013



University of Economics

Prague

Epistemologie dat v současné vědě

Jindřich Mynarz



Abstract

Contemporary science is dominated by positivist epistemology of data, which builds on metaphysical realism and the ideal of mechanical objectivity, yet suffers from a number of flaws. Shortcomings of this approach were identified in many critical responses and led to a new problematization of the established concept of data. The often criticised aspects of this concept remark on data being embedded in the context of its making, and point out the mediation of data and its openness to manipulation. In recent years, the function of data gained an unprecedented importance due to the rising demands of science for data, which attracted attention to this formerly unproblematic concept. Several alternative approaches to epistemology of data appeared, of which this text introduces the positions proceeding from constructivism and rhetoric. The presented paper draws heavily on critical literature in epistemology of data. Due to its summarizing character, it may be understood as a synthesis and reconfiguration of the existing thoughts on the topic. In this way, the paper offers a contribution to rhetorical argumentation in the discourse of data in contemporary science.

Keywords: data, epistemology, constructivism, rhetoric

Abstrakt

Dominantní pozitivistická epistemologie dat v současné vědě, která staví na metafyzickém realismu a ideálu mechanické objektivit, trpí řadou trhlin. Tyto nedostatky byly identifikovány v mnohých kritických ohlasech, které vedou k nové problematizaci ustáleného konceptu dat. Mezi kritizovanými aspekty tohoto konceptu se opakují varování ohledně zasazenosti dat do kontextu jejich tvorby, jejich zprostředkovanost nebo otevřenost manipulacím. Funkce dat v posledních letech nabývá na důležitosti kvůli stoupající datové náročnosti vědy, a proto je dříve neproblematickému konceptu dat věnováno více pozornosti. Objevují se také alternativní přístupy k epistemologii dat, z nichž tato práce představuje směr konstruktivistický a rétorický. Prezentovaný text lze díky své kompilativní povaze, čerpající z kritické literatury zabývající se epistemologií dat, pojmout jako syntézu a rekonfiguraci existujících myšlenek k danému tématu. Je možné ho rovněž přijmout jako jeden z příspěvků k rétorické argumentaci v diskurzu dat současné vědy.

Klíčová slova: data, epistemologie, konstruktivismus, rétorika

Úvod

Navzdory tomu, že etymologie slov mnohdy nijak nesouvisí s jejich způsobem užití, napovídají kořeny slova ‘data’ mnoho o původech toho, jakým způsobem je tento pojem používán. Data, jak popisuje Rosenberg [18, s. 18], pocházejí z plurálu latinského výrazu ‘datum’, který je minulým přičestím středního rodu slovesa ‘dare’, jenž se do češtiny překládá jako ‘dát’. ‘Datum’ pak lze přeložit jako ‘něco daného’. V souladu s tímto významem jsou pak data v obecném užití obvykle brána jako daná; jako něco, co není třeba zpochybňovat.

V protikladu k této interpretaci se staví konstruktivistická epistemologie, podle níž nic není dáno, jelikož vše je produktem lidské konstrukce. Bachelard tak například píše:

“Pro vědeckou mysl je vše vědění odpovědí na otázku. Pokud nejsou žádné otázky, pak žádné vědecké znalosti nemohou existovat. Nic není zjevné. Nic není dáno. Vše je zkonstruováno.” [3, s. 25]

Z rozcházejícího se chápání dat pak vychází soudobá kritika, která podrývá zavedený status dat ve vědě. Na jedné straně jsou data přijímána jako přímý obraz skutečnosti, zatímco na druhé straně jsou brána jako artefakt lidské tvorby. Tyto rozporné přístupy se promítají do problematizace konceptu dat a přináší mnoho pochybností. Například Poovey ve svém textu věnovaném historii moderního faktu formuluje řadu otázek, které lze stejně dobře aplikovat na koncept dat:

“Co jsou fakta? Jsou to nevyvratitelná data, která jednoduše ukazují pravdu? Nebo jsou to kousky důkazů shromážděné za účelem přesvědčení ostatních o vysvětlované teorii? Existují fakta ve skutečnosti a podobně jako kamínky čekají na toho, kdo je sebere? Nebo jsou zhotovena a proto prodchnuta všemi společenskými a osobními faktory, které se podílí na každém aktu lidské tvorby? Jsou fakta mimo rámec interpretace? Nebo jsou právě součástí výkladu a jeho příznačným ztělesněním namísto bodu, kde je výklad zahájen?” [37, s. 1]

Prezentace možných odpovědí na tyto otázky je předmětem tohoto textu. Na základě cest, jimiž se chápání dat ve vědě vyvíjelo, je dále představena kritika dominantní realistické epistemologie dat a jsou uvedeny alternativní epistemologie, které čerpají z konstruktivismu a rétoriky.

Stručná historie pojmu ‘data’

Pojem ‘data’ je v užití již dlouho, avšak jeho soudobý význam souvisí až s nástupem modernity [18, s. 15]. Jedno z prvních dochovaných užití tohoto výrazu se objevuje v Eukleidově knize nazvané právě *Data* [16]. Jde o text popisující metody řešení a analýzy problémů, v němž data figurují jako to, co je s ohledem na hypotézu

známo, nebo to, co může být demonstrováno jako známo. Data zde slouží jako daná východiska, z nichž mohou být vyvozeny nové znalosti.

Použití dat pro označení daností zůstalo v užití přinejmenším do 17. století. V matematice, filosofii a teologii ‘data’ označovala daná východiska, která nejsou zpochybňována [18, s. 19]. V teologii byl tento výraz uplatňován například pro věci dané bohem nebo Biblií.

Ke svému modernímu způsobu užití se pojem začal přiklánět začátkem 18. století. Namísto významu nezpochybnitelných daností se objevilo užití dat ve smyslu výstupů experimentů, zkušeností nebo sběru [18, s. 36]. O prosazení tohoto přístupu do obecného povědomí se postaral především pozitivismus. Z epistemologického hlediska lze toto moderní pojetí označit jako metafyzický realismus dat.

Metafyzický realismus dat

V souladu s metafyzickým realismem lze data z objektivně existující skutečnosti ‘sbírat’. V rámci realismu data nabývají role přesného záznamu nebo věrné reprezentace skutečnosti. Rétorika vědeckých ‘objevů’ předpokládá, že jimi odhalené znalosti již ve skutečnosti existují a úkolem vědy je tyto znalosti najít. Prototypickým příkladem přesného zachycení skutečnosti je fotografie. Fotografie jsou *“surové reprezentace přirozeného světa,”* které představují *“unikátní a doslovné transkripce přírody - druh vědeckého záznamu”* [18, s. 4, příl.].

Ve vědě jsou data obecně chápána jako záznamy předem navrženého strukturovaného pozorování, které se řídí daným protokolem [20]. Sběr dat je charakterizován jako pozorování bez zásahu do pozorované skutečnosti. Žádoucí objektivita předpokládá oddělení vnímajícího subjektu od pozorované skutečnosti. Podle Von Glasersfelda [44, s. 2] pak metafyzický realismus *“tvrdí, že jako ‘pravdivé’ můžeme označit jen to, co koresponduje s nezávislou, ‘objektivní’ realitou.”* V tomto myšlenkovém rámci mají data privilegovanou úlohu před subjektivním popisem vzdáleným ‘pravé’ skutečnosti.

Epistemologické privilegium dat

Důraz na data se ve vědě prosazuje zejména v posledních letech, kdy se velké objemy dat stávají široce dostupnými. Data samozřejmě hrála ve výzkumu podstatnou roli i dříve; např. Nelson [32] uvádí Rudolfské tabulky Johanna Keplera jako raný příklad vědeckého užití dat. Nicméně jejich příznačná role byla datům *“dána až v epistemologii, kterou spojujeme s modernitou”* [37, s. 1].

Kvantifikace v podobě dat začala v moderní vědě s postupem času čím dále více *“přispívat k dojmu objektivitu vědecké literatury”* [19, s. 37]. Průměrný počet tabulek s daty použitých ve vědeckých článcích se mezi 19. a 20. stoletím takřka zdvojnásobil. Ve vzorku článků z 20. století zhruba polovina obsahovala tabulku, přičemž mezi nimi byl průměrný počet 5 tabulek na článek [19, s. 182]. Ve 20. století se preference vědy výrazně přichýlily ke kvantitativním faktům namísto faktů kvalitativních. V jazyce vědy se tento příklon promítá do rozlišení ‘tvrdých dat’,

jinž jejich kvantitativní povaha dodává auru nezpochybnitelnosti, a ‘měkkých dat’, jenž lze díky své kvalitativní formě takřka libovolně ohýbat. V některých případech toto upřednostňování zachází až k situacím, kdy je kvantitativním datům *“přisuzována značná důležitost, ačkoli jejich platnost nikdo nehájí s opravdovým přesvědčením”* [38, s. 8]. Valná většina vědců ve svých pracích poskytuje mechanická nebo matematická vysvětlení pro popisovaná fakta a jejich *“argumentační strategie pro ustanovení faktů a vysvětlení se obvykle točí okolo porovnání dat”* [19, s. 188]. Matematická vysvětlení s odkazy na data jsou privilegována pro svou domnělou eleganci a přehlednost [21]. Koncem 20. století se pak objevuje výrazná tendence upřednostňovat *“porovnávání velkých souborů dat, přičemž matematické metody jsou aplikovány kdykoli je to možné”* [19, s. 231]. Postupně tak dochází k *“překotné komodifikaci dat,”* kdy jsou data prezentována jako *“kompletní, zaměnitelné produkty ve snadno směnitelných formátech,”* což však může vést k *“podporování dezinterpretace, spoléhání se na slabé nebo zavádějící zdroje, případně preferování dostupnosti dat před jejich kvalitou”* [15, s. 7].

Věda řízená daty

V nedávných letech se však důraz na užití dat ve vědě zvýšil natolik, že někteří mluví o nástupu nového metodologického paradigmatu *vědy řízené daty* [27]. Tento přístup bývá označován jako *čtvrté paradigma vědy*, které prostřednictvím datové intenzivního výzkumu, v němž počítače napomáhají k nacházení znalostí v datech, rozšiřuje předešlá tři paradigma; a to paradigma empirického pozorování, paradigma tvorby vysvětlujících modelů a paradigma simulace pro vhled do složitých jevů [33]. Data lze považovat za produkt kvantitativního výzkumu, který je obzvláště privilegován pro roli vědeckých důkazů.

Extrémní případy prosazování paradigmatu vědy řízené daty si vysloužily označení ‘datový fundamentalismus’ [11]. Například Anderson [1] ve svém kontroverzním článku z roku 2008 provolává nástup velkých dat jakožto *“konec teorií”* a píše, že *“čísla mluví za sebe”* a hypotézy se tak stávají nepotřebnými. Jenže, jak upozorňuje Keller [24, s. 130], *“problém s tímto argumentem je samozřejmě to, že data nikdy nemluví sama za sebe.”* Bez ohledu na tyto výtky však zbavení se nutnosti formování hypotéz podle některých autorů vede k ‘očistění’ vědy:

“Ve světě malých dat, v němž bylo dostupných dat velmi málo, jak kauzální šetření, tak korelační analýza začínala s hypotézou, která byla následnými testy falsifikována nebo verifikována. Avšak protože obě metody vyžadují začít s hypotézou, jsou stejně náchylné vlivu předsudků a chybné intuice.” [30]

Důvodem, proč tvorba hypotéz ztrácí v tomto paradigmatu vědy na důležitosti, může být ‘bezduševní účinnost dat’. Podle některých autorů [21] totiž jednoduché modely (resp. hypotézy) opatřené dostatkem dat nevyhnutelně překonávají složité modely trpící nouzí o data. Ve své krajní podobě věda založená na datech převrací obvyklý vztah mezi hypotézou a daty, v němž má hypotéza primární roli a data

slouží k jejímu potvrzení nebo naopak falsifikaci. Namísto toho toto paradigma prosazuje postupy, pomocí nichž lze data indukci zobecnit na platné hypotézy. Výzkumem těchto postupů se zabývá oblast dolování z dat, kde výsledkem převodu jednotlivého v datech na obecné může být například sada asociačních pravidel.

Ve vědě založené na datech jsou hypotézy svou podstatou považovány nedůvěryhodné, pokud jejich spolehlivost není podložena daty. K tomu však kritický racionalismus Karla Poppera připomíná, že data podporující hypotézu jí neverifikují, jen falsifikují s daty nekompatibilní hypotézy.

Důvěryhodnost připisovanou datům lze ilustrovat populárním úslovím *“V boha věříme, ostatní ať přinesou data.”* Data tak v současné vědě fungují jako doklad svědčící o pravdivosti prezentovaných tvrzení a například Markham [29] zmiňuje dojem ‘instantní důvěryhodnosti’, který data vyvolávají.

Naproti tomu Boyd a Crawford [9, s. 663] tento mnohdy nekriticky přijímaný přístup popisují jako *“široce rozšířené přesvědčení, že rozsáhlá data poskytují vyšší formu intelligence a znalostí, která může nabídnout dříve nedosažitelný vhled obdařený auroou pravdy, objektivitu a přesnost.”* K tomu například Keil dodává, že *“věda založená na datech je selháním představivosti”* [23]. Keil zdůrazňuje, že ve vědě nelze ignorovat tvorbu hypotéz, které nabývají podob modelů nebo teorií. Naopak je nezbytné kombinovat jak empirická pozorování, tak tvorbu hypotéz. Velký objem dat nepomůže, pokud není konfrontován s užitečnou teorií. Zvýšením objemu dat sice vzroste ‘podpora’ prevalentních hypotéz, ale souběžně s tím obdobně naroste šum v datech. Proto může velký objem navzdory obvyklým očekáváním také vést ke znásobení problémů v datech. Privilegovanost dat lze osvětlit mýtem jejich mechanické objektivitu.

Mýtus mechanické objektivitu dat

Epistemologické privilegium dat pramení z ideálu mechanické objektivitu dat [18], který ignoruje kontextualitu, zprostředkovanost a manipulovatelnost dat. Domnělou absencí lidského vstupu (např. při fotografování) a minimalizací nežádoucích vlivů tak údajně lze docílit objektivitu. Víra v neutrálnost, autonomii a objektivitu dat je velmi rozšířená. Porter například tvrdí, že *“když filosofové mluví o objektivitě vědy, tak obecně míní její schopnost znát věci, jaké ve skutečnosti jsou”* [38, s. 3]. V souladu s tímto požadavkem metafyzický realismus považuje data za přímou reprezentaci skutečnosti. Data však přesným obrazem komplexní skutečnosti nemohou být, protože jí nevyhnutelně redukuje a zanedbávají tak detaily, které jsou považovány za nepodstatné pro účel, pro nějž byla data vytvořena. Vysokou úroveň redukce data ztrácí schopnost reprezentovat skutečnost, a proto mohou být považována nanejvýše za její aproximaci [29]. Příčin selhání datových reprezentací skutečnosti lze identifikovat více a jsou dále rozebrány v následujících sekcích textu.

Je však třeba zmínit, že existují i jiné způsoby formulace objektivitu vědy. Může jím být vlivná definice objektivitu jako *“schopnosti dosáhnout konsenzu”* [38, s. 3], nebo jako *“poctivosti a nestrannosti”* [38, s. 4]. Ideál mechanické objektivitu dat

je však nedosažitelný, protože data jsou nutně zprostředkována a jejich tvorba je vždy zasazena do kontextu, kterému se nelze vyhnout a ani jej nelze reprodukovat.

Kontextualita dat

Na podobě dat se významnou měrou podílí kontext, v němž jsou vytvořena. Význam dat ve vědě je *“striktně závislý přinejmenším na přesném porozumění toho, jak, kde a kdy byla data vytvořena”* [15]. *“Produkce znalostí není nikdy oddělena od producenta znalostí”* [5] a ani data nelze získat bez přímého či nepřímého lidského vstupu, takže na nich má lidské myšlení vždy podíl. Do přímých vjemů se tak mohou dostat mentální obsahy vnímajícího a do nepřímého vnímání skrze přístroje se může promítnout pohled jejich tvůrců, který do nástrojů zabudovali. Bachelard toto ilustruje slovy [3, s. 24]: *“Když přemýšlíme o skutečnosti, tak to, co velmi dobře známe, vrhá stín na to, co bychom měli vědět.”* Je proto zapotřebí mít na paměti *“situované, materiální podmínky produkce znalostí”* [18, s. 4], kvůli nimž je výsledná data *“třeba chápat jako zarámovaná a rámuující”* [18, s. 5].

Kromě samotných tvůrců dat se na jejich zarámování zásadním způsobem podílí prostředí, v němž jsou vytvářena. Na tento vliv upozorňuje například Magee:

“Znalostní systémy jsou až příliš často charakterizovány výrazy esencialismu - jakoby, jak etymologie slova ‘data’ napovídá, byly pouze pláštěm pro neutrální empirické danosti. [...] naopak, tyto systémy s sebou vždy nesou předpoklady kultur, které je navrhují a používají, a které jsou tak v širším slova smyslu za ně zodpovědné.” [28, s. 15]

Prostředí může stanovovat způsob získávání dat, například standardizací rozdílných metod, které mohou s časem evolvovat. V průběhu let tak může kupříkladu dojít k reklasifikacím v rámci systémů kategorií, které zásadně zhoršují porovnatelnost dat v čase [13].¹

Vliv kontextu na data však nelze eliminovat a stejně tak jej nelze reprodukovat. Díky své diskrétní povaze, kdy *“každý prvek dat je individuální, oddělený a oddělitelný, přičemž je ale stále podobný ostatním prvkům”* [18, s. 8], data lze částečně dekontextualizovat, avšak snaha o úplné odstranění vlivů kontextu nemůže být nikdy úspěšná.

Někteří si slibují, že očištění dat od zkreslujících vlivů kontextu a subjektivity je možné dosáhnout, pokud jsou data k dispozici ve velkém objemu a obsahují mnoho variací záznamů téhož fenoménu. Příslib velkých dat spočívá v tom, že jejich kombinací a agregací je lze neutralizovat, a snížit tak jejich individuální zabarvení, čímž se data přiblíží objektivní skutečnosti. Selháním tohoto přístupu je opomenutí arbitrárnosti zvoleného způsobu agregace a ignorování nekompletnosti dat. Volba způsobu agregace je subjektivním aktem, při němž použití arbitrárních konceptualizací (např. systémů kategorií) znamená, že výsledná agregovaná data se nakonec mohou skutečnosti ještě více vzdálit. Jakkoli jsou data rozsáhlá,

¹Tento problém se týká například pomíjivých skupin, jako je skupina 1 % nejbohatších, kdy díky proměnlivosti této kategorie nelze srovnávat různé časové snímky dat o ní vypovídajících.

stále zůstávají pouze nekompletním vzorkem, jehož výběr může vynechat to, co je důležité. Rozsáhlost dat nezaručuje jejich reprezentativnost, jelikož data jsou předmětem omezení a předsudků nezávisle na své velikosti. Stejně tak kvantita nemůže zastoupit kvalitu a konzistenci dat, která se díky vlivu proměnných kontextů může rozcházet. V případě vzorků dat nejsou absolutní hodnoty nikdy přesné a relativní hodnoty jsou zatíženy tím, co do vzorku promítá jeho způsob výběru a agregace.

Snaha o dekontextualizaci dat může být navíc škodlivá pro kontext jejich užití. Boyd a Crawford připomínají, že pokud jsou *“data z kontextu vytržena, ztrácí svůj význam a hodnotu”* [9, s. 670]. Data představují médium, jehož užití předpokládá aktivní účast a porozumění uživatele. Vědění není pasivní proces [44, s. 9]. Ačkoli předpoklad objektivně existující skutečnosti vytváří společný referent dat, na nějž se váže společné pochopení, tak sdílené porozumění dat zůstává převážně fikcí, jelikož jejich interpretace je vždy závislá nejen na referentu, ale také na kontextu [29].

Mechanická objektivita se pokouší kontextuální vlivy redukovat do přesně vymezeného protokolu. Tvorba dat je proto podřizována striktním pravidlům. Mechanická objektivita je tak dána schopností následovat pravidla a řídit se stanoveným protokolem [38, s. 4]. Rolí protokolu je nastavení kontrolovaného kontextu a minimalizace nežádoucích vlivů, které by se mohly projevit na vytvářených datech. Transparentní a zdokumentovaný protokol přípravy dat, který obsahuje detailní informace o původu dat, slouží pro zvýšení jejich důvěryhodnosti. Bird tvrdí, že to, *“co dělá něco předmětem znalostí plynoucích z pozorování je spolehlivost a nesporná povaha mechanismu, který tuto věc vytváří”* [4, s. 10]. Uživatelé dat tak pomocí vyhodnocení *“adekvátnosti experimentálních podmínek, v nichž byla data vytvořena,”* [27] mohou určit, jakou spolehlivost lze od dat čekat a jaká je jejich evidenční hodnota.

Zároveň omezení protokolem usiluje o možnost reprodukce dat. Leonelli [26] nicméně uvádí, že *“data jsou idiosynkratická pro konkrétní experimentální kontexty a typicky se nemohou objevit mimo tyto kontexty.”* Díky nevyhnutelnému zasazení do neopakovatelného kontextu tedy data nelze v úplnosti reprodukovat. Nanejvýše se lze pokusit o reprodukci metod, pomocí nichž byla data vytvořena, díky čemuž je možné dospět k jiným, ale částečně kompatibilním datům.

Zprostředkovanost dat

Bezprostřednost přisuzovaná datům vychází z touhy po přímém poznání skutečnosti. ‘Surovým’ datům je připisována kvalita primárnosti. Jde o data ‘přímo’ od zdroje, jímž je skutečnost. Tato domnělá kvalita dat souvisí se zdánlivou přirozeností jejich mechanické tvorby. Lze tak podlehnout dojmu, že hodnota dat je závislá na přímočarosti jejich odvození ze skutečnosti. Například data z automatizovaných senzorů tak můžou být vnímána jako podstatně důvěryhodnější než výpočty impakt faktoru, které jsou založené na nepřímě vyvozených podkladech výrazně vzdálených skutečnosti. Díky své předpokládané bezprostřednosti jsou data často chápána v souladu se svou etymologií jako danosti axiomatické povahy,

kteří není třeba zpochybňovat [20]. *“Na první pohled data zjevně předchází znalosti: představují výchozí bod pro to, co víme, kdo jsme a jak komunikujeme”* [18, s. 2]. *“O datech se nediskutuje,”* píše Markham [29], protože jsou chápána jako popis toho, co předchází vlastní argumenty. Samotná data se tak v tomto pojetí vyhýbají interpretaci, analýze a tím také subjektivnímu zabarvení. Nicméně, jak připomínají Boyd a Crawford [9, s. 667], *“tvrzení o objektivitě jsou nutně vytvářena subjekty na základě subjektivního pozorování a výběru.”*

Předpoklad preanalytické povahy dat však byl podroben kritice, díky níž došlo k problematizaci konceptu ‘surových dat’ a platnost tohoto předpokladu byla zpochybněna. Dewey tuto domněnku kritizuje již v roce 1929:

“[...] všechno soupeření a s ním související problémy vyrůstají ze stejného kořene. Vycházejí z předpokladu, že pravdivým a validním objektem znalostí je to, co předchází a je nezávislé na fungování vědění. Vycházejí z nauky, že znalost je pochopení nebo pohled na skutečnost, aniž by bylo provedeno cokoli, co by změnilo její předcházející stav - což je učení, které je zdrojem předpokladu oddělenosti znalostí od praktických aktivit.” [12, s. 196]

Podobný kritický přístup k datům považuje Keller již v roce 1985 za všeobecně osvojený, když píše o tom, že *“v současnosti je takřka truismem, že nic jako surová data neexistuje; všechna data předpokládají interpretaci”* [24, s. 130]. Bowker k tomu poznamenává, že *“surová data jsou současně oxymorónem a špatným nápadem; naopak, data by měla být pečlivě připravována”* [8, s. 184].

Klíčovým bodem těchto kritik je uznání toho, že interpretace je obsažena již ve samotném pozorování při vlastní tvorbě dat. Podle Nunberga jsou vlastnosti, které připisujeme informacím, a které lze vztáhnout i na data - jako jejich *“přenositelnost, kvantifikovaná povaha, interpretační transparentnost nebo autonomie - jednoduše reifikací různých principů interpretace, které používáme při čtení informací”* [34]. Data a *“čísla jsou interpretována, protože v sobě zahrnují teoretická východiska pro to, co má být počítáno, jak máme rozumět materiální realitě a jak kvantifikace přispívá k systematickým znalostem o světě* [37, s. xii]. Data tedy nelze brát za *“jednoduchá pozorování jednotlivin, která by byla imunní zájmům a teoretickým dohadům všeho druhu”* [37, s. xxiv].

Data jsou z mnoha důvodů vždy zprostředkována, a tak, jak píše Bachelard [3, s. 24]:

“Poznání skutečnosti je světlo, které vždy na některé koutky a skuliny vrhá stín. Není nikdy bezprostřední a nikdy úplné”

K tvorbě dat dochází vždy skrze média. Příkladem takových médií mohou být nástroje, jako je mikroskop, nebo modely, které mohou data syntetizovat. Díky médiím může věda vypovídat i o objektech a vlastnostech, které přímo nelze pozorovat, jako jsou například vzdálené, dávno zaniklé galaxie, které lze zaznamenat skrze teleskopy [6]. V obecném pojetí *“vědecké teorie predikují a vysvětlují fakta o*

‘pozorovatelných’ objektech a vlastnostech, které mohou být vnímány smysly, někdy posílenými pomocí nástrojů” [6, s. 303].

Při svém vzniku data nezbytně procházejí skrze modely skutečnosti. Model je médiem poznání. *“Racionální vnímání světa je založeno na idealizovaných modelech, které jsou pouze aproximací reality”* [25, s. 220], nicméně *“náš pohled může být dostatečně podobný pro většinu účelů, takže se skutečnost může jevit objektivní a stálá”* [25, s. 228]. Přesto se však objevují hlasy, které prosazují *“představu, že ‘spolehlivá věda’ musí znamenat ‘nevyvratitelný důkaz daty z pozorování’, zatímco modely jsou neodmyslitelně nedůvěryhodné”* [14, s. xviii]. *“Ať data mluví sama za sebe,”* [24, s. 129] dožadují se někteří vědci, kteří volají po surových, nezprostředkovaných datech. Tento předpoklad, že se bezprostřednosti lze dobrat *“čekáním na data nezávislá na modelech”*, označuje Edwards za zcestný [14, s. xiii]. Jak píše dále, tak v současnosti se *“žádný sběr signálů nebo pozorování - dokonce i ze satelitů, které mohou ‘vidět’ celou planetu - nestává globálním v čase a prostoru, aniž by nejprve neprošel řadou datových modelů”* [14, s. xiii]. Závislost dat na modelech lze ilustrovat na příkladu předpovědi počasí a predikcí změn klimatu, v nichž *“pouze 10 % dat používaných v globálních predikčních modelech pochází ze samotných odečtů ze senzorů. Zbývajících 90 % dat je syntetizováno dalšími počítačovými modely.”* [14, s. 21] Data jsou stejně jako modely a teorie nedokonalé aproximace skutečnosti. Přesto však, podobně jako Box [7, s. 424] tvrdí, že *“všechny modely jsou špatné, ale některé jsou užitečné”*, tak analogické tvrzení lze formulovat pro data.

Manipulace daty

Díky zprostředkovanosti je možné data manipulovat a účelově rekonstruovat. Zkreslení ovlivňující data mohou být jak vědomá, tak nevědomá. Vliv kontextu na data se může podepsat na sotva patrných pokřiveních vytvořených dat. Jak upozorňuje Fanelli [17], *“vědecké výsledky mohou být zkresleny mnoha způsoby, které mohou být mnohdy natolik jemné, že uniknou pozornosti vědců.”* Navzdory tomu, že s vědou je obvykle asociována *“poctivost a nestrannost”* [38, s. 4], tak významný podíl zkreslení dat je vědomý. Před rizikem manipulací ve vědě varuje již roku 1830 Charles Babbage:

“O manipulacích. Toto je um mnoha forem, jejichž účelem je obdařit běžná pozorování zdáním a charakterem nejvyšší úrovně přesnosti. Jedním z početných manipulativních procesů je zhotovení velkého množství pozorování a vybrání pouze těch, které souhlasí nebo skoro souhlasí. Pokud je zaznamenáno 100 pozorování, pak by manipulátor musel mít velkou smůlu, kdyby z nich nemohl vybrat 15 nebo 20, které postačí pro servírování.” [2, s. 178]

V anonymních průzkumech se k manipulaci s daty přiznalo zhruba 2 % vědců. Přibližně třetina účastníků průzkumů připustila použití pochybných vědeckých praktik. Avšak vzhledem k tomu, že zde jde o citlivé téma, tyto odhady jsou

pravděpodobně konzervativní [17]. Kromě záměru totiž může být za zkreslováním dat také lenost a pochybení.

Mezi úmyslnými zkresleními dat lze uvést [31] například ignorování nepříhodných dat, data odpovídající na sugestivní otázky, přílišné zobecnění, zkreslený vzorek (např. nenáhodný), špatný odhad chybovosti, neoprávněný předpoklad kauzality, manipulaci dat výběrem nebo fabrikací nebo nalézání statisticky nevýznamných korelací ve velkých objemech dat. Například pro zamlžení a zhoršení srozumitelnosti dat může být použita redukce na agregace [13].

Alternativní epistemologie dat

Namísto metafyzického realismu lze na epistemologii dat nahlížet z alternativních pohledů, které netrpí mnohými z výše popsanych neduhů. Tento v zásadě *“pozitivistický obraz struktury vědeckých teorií je nyní všeobecně odmítán”* [6, s. 304] a jeho místo obsadily přístupy, které lze označit za postmoderní, ale které mnohdy čerpají ze staršího myšlení a sahají například až k rétorickým počátkům filosofie. V tomto textu jsou představeny přístupy konstruktivistické epistemologie a rétoriky, které jsou vzájemně kompatibilní.

Konstruktivistická epistemologie dat

Konstruktivistická epistemologie je založena na předpokladu, že veškeré znalosti jsou konstrukcí člověka. S metafyzickým realismem se tento směr zásadně rozchází v tom, že nevyžaduje koncept objektivně existující skutečnosti. Chápání konstruktivismu výhradně jako odmítnutí konceptu objektivní reality by však bylo příliš zjednodušené. Konstruktivismus převrací vztah mezi daty a skutečností tvrzením, že data jimi popisovanou skutečnost konstituují. Podle konstruktivistického přístupu tak *“data nejsou nacházena, ale vytvářena”* [20].

Některé z ústředních tezí konstruktivistické epistemologie lze dobře vidět již ve formulacích Giambattisty Vica z 18. století. Ve spisech tohoto intelektuálního předchůce konstruktivismu lze nalézt tvrzení, že *“věda je znalostí původu a způsobů, jimiž jsou věci tvořeny”* a proto *“můžeme znát jen to, co sami konstruuje”* [44, s. 7]. V tomto přiznání spočívá jeden z rozdílů mezi vědeckou a předvědeckou myslí, protože *“zatímco předvědecká mysl skutečnost vlastní, vědecká mysl ji konstruuje a rekonstruuje, přičemž se sama neustále mění”* [3, s. 9].

Velký podíl na konstruktivistické epistemologii může mít příklon filosofie k jazyku. Očima konstruktivismu lze číst například dílo antropologa a lingvisty Edwarda Sapira, v jehož pojetí je ‘svět’ konstruován jazykem skupiny:

“Ukazuje se, že ‘skutečný svět’ je velkou měrou nevědomě vystaven na jazykových zvycích skupiny. Žádné dva jazyky nejsou dostatečně vzájemně podobné, abychom mohli uvažovat, že reprezentují stejnou sociální realitu. Světy, v nichž různé společnosti žijí, jsou odlišné světy, a ne jeden stejný svět s rozdílným označením.” [39, s. 221]

Homomorfismus mezi daty a skutečností, kdy data odpovídají zkušenosti skutečnosti, tedy pro konstruktivismus není zapotřebí. Namísto toho jsou data a znalosti tím, co na skutečnost sedí a konzistentně s ní funguje. Pro ilustraci tohoto vztahu Von Glasersdorf nabízí příměr klíče, který zapadá do zámku, podobně jako data, která pasují na skutečnost [44, s. 3].

“Tvzení, že věda je společensky konstruována byla až příliš často vnímáno jako útok na její platnost a pravdivost” [38, s. 11], proto však konstruktivismus nabízí kritéria pravdivosti nahradit konceptem vnitřní konzistence a bezrozpornosti v systému znalostí [44, s. 9]. Za těchto podmínek si výzkum lze představit jako generativní proces, který vytváří data a eliminuje nefunkční znalosti. Příkladem znalostí, které jsou odhaleny jako nekonzistentní, mohou být následky ‘apophenie’, což je jev *“vidění vzorů tam, kde žádné neexistují”* [9, s. 668]. Apophenií mohou být postiženi například datoví analytici, když podlehnou dojmu, že rozpoznali kauzalitu v datech, zatímco jde o idiosynkratickou konstrukci pozorovatele.

Konstruktivistickému přístupu přispívá to, že tvorba dat je vždy do jisté míry aktem klasifikace. Díky své poddajnosti mohou být data odlita pomocí zvolených datových struktur a konceptualizací. Jakmile je však klasifikace zavedena, stává se součástí dat a je obtížné ji odlišit. Na arbitrárnost datových struktur přivádí pozornost Kent, který píše:

“Datové struktury jsou umělé formalismy. Od informací se liší podobně jako gramatiky nepopisují jazyky, které používáme, a formální logické systémy nepopisují, jak přemýšlíme.” [25, s. xix]

Podobně jako je jazyk společenství používán ke konstrukci sdíleného světa, tak datové struktury jsou podkladem vytvářejícím sdílené porozumění datům. A stejně tak jako *“různé druhy map, každý druh struktury má své silné stránky a slabiny, které slouží rozdílným účelům a zamlouvá se různým skupinám lidí v různých situacích”* [25, s. xix].

Ačkoli na různé stránky konstruktivismu odkazuje kritika metafyzického realismu, částečně zmíněná v sekci , tak jeho epistemologie v mnoha oblastech vědy nedominuje. Zatímco například pro pole psychologie Piaget v roce 1980 píše, že *“padesát let zkušeností nás naučilo, že znalosti nejsou výsledkem záznamů pozorování bez strukturující aktivity na straně subjektu”* [36, s. 377], a principy konstruktivistické epistemologie jsou v humanitních a společenských vědách dobře přijímány, tak zato v oblasti přírodních věd jsou spíše ignorovány [22] a převládají v nich pozůstatky metafyzického realismu.

Rétorika dat

Rétorika nabízí pojetí dat, které lze přijmout jako komplementární k chápání konstruktivistické epistemologie. Tato slučitelnost je vidět zejména v případě ontologického přístupu k epistemickému chápání rétoriky, který rozlišuje Brummett [10]. Ontologický výklad rétorické epistemologie prohlašuje, že *“diskurz neslouží pouze*

k odhalení pravdy nebo učinění pravdy účinnou. Namísto tvorby pravd o skutečnostech tento diskurz skutečnosti vytváří" [10]. Funkcí rétoriky tedy není jen přesvědčování a tvorba ospravedlnění, ale i tvorba samotných obhajovaných tvrzení. Scott jako jeden z prvních, kdo propojil rétoriku s epistemologií, píše, že *"rétorika může být tak nahlížena nikoli jako činnost, které dělá pravdu účinnou, ale jako otázku tvorby pravdy"* [41, s. 13]. K epistemologickému konstruktivismu směřuje i Scottova poznámka, že *"'pravda', může být samozřejmě pochopena v různých smyslech. Pokud ji bereme jako apriorní a neměnnou, pak pro rétoriku nenajdeme jiné užití než v promluvách ke slabším."* [41, s. 9]

Jak nasvědčuje užití pojmu 'data' ve spisech Eukleidových [16], tak již tehdy byl tento koncept uplatňován rétorickým způsobem. Data jsou dle svého etymologického výkladu *"tím, co je 'dáno před argumentem', co je dáno pro řečnický základ"* [18, s. 7]. Tvorbu dat ve vědě lze zasadit do diskurzu, v jehož rámci probíhá rétorická argumentace. Data jsou konstruována jako jeden z produktů tohoto diskurzu, v němž slouží primárně přesvědčovací funkci. Výběr a zpracování dat lze přizpůsobit tak, aby data posloužila žádanému účelu v argumentaci. Jestliže jsou tvrzení napadána, je na jejich autorech, aby je ospravedlnili. *"Pokud jsme vyzváni, je na nás, abychom vytvořili jakákoli data, fakta nebo jiné podklady, které považujeme za relevantní a dostatečné pro ospravedlnění počátečního tvrzení"* [43, s. 13]. Produkci dat tak lze považovat za specifický řečový akt, který lze použít v argumentaci pro ospravedlnění předchozích nebo nadcházejících tvrzení.

Rétorický argument představuje alternativu k analytické logice. Podobně jako v logice, tak v rétorice, *"jsou-li dána určitá data, lze vyvodit určité úsudky"* [18, s. 7]. Data však nepatří do rámce analytické logiky, protože nemají žádnou pravdivostní hodnotu. *"Pokud je prokázáno, že fakt neplatí, pak přestává být faktem. Neplatná data avšak zůstávají daty"* [18, s. 18]. Užití dat je tedy rétorické. Gitelman toto rozlišení shrnuje slovy *"fakta jsou ontologická, důkazy jsou epistemologické a data jsou rétorická"* [18, s. 18].

Rétorika má však ve vědě špatnou reputaci. Na nebezpečí rétoriky upozorňuje například Thomas Sprat roku 1667, kdy vydává své pojednání o historii britské Královské společnosti: *"A aby toho dosáhli, usilují o oddělení vědění od Přírody, od zabarvení Rétoriky, způsobů Rozmaru nebo nádherných klamů Pověr"* [42, s. 62]. Historicky je rétorika spojena s úmyslnými, manipulativními užitími dat. Některým z těchto užití se věnuje předcházející sekce. Manipulaci dat, například za účelem získání grantu, lze pokládat za druh rétorické argumentace. Příklady užití dat pro rétorické účely můžeme najít například v propagandistických infografikách,¹ debatách ohledně existence globálního oteplování nebo v datech z předvolebních průzkumů, jejichž tvůrci jsou často obviňováni ze záměrné manipulace.

V případě dat je to jejich možnost agregace, která *"jim propůjčuje jejich potenciální sílu a jejich rétorickou váhu"* [18, s. 8]. Příkladem rétorického vytvoření dat je rekonceptualizace novin jako databáze provedená Angelinou Grimké Weld, jejím manželem Theodorem a její sestrou Sarah [18, s. 90]. V tomto případě byla zkom-

¹Jak zmiňuje Gitelman, *"vizualizace dat zesiluje jejich rétorickou funkci"* [18, s. 12].

pilována data o otroctví, založená například na novinových inzerátech na zběhlé otroky. Tato data byla rekontextualizována jako svědectví brutality otrokářů tím, že jejich vlastní slova obrátila proti nim.

Moderní rétorika má však podstatně širší záběr než jen manipulace a přesvědčování. Ontologický přístup zmiňovaný Brummettem rétoriku staví do role dimenze přítomné ve všem epistemologických aktivitách [10]. Rétorická dimenze se podepisuje i na datech, které užívá věda. Ačkoli se objevují prohlášení, že *“data jsou apolitická”* [35], tak data nikdy nejsou nestranná a vždy je nutné počítat s tím, že obsahují skrytou rétorickou agendu. I když praktické analýze dat ve vědě povětšinou vědomý rétorický přístup chybí [40], vzrůstající objem výzkumné práce v této oblasti¹ dává tušit alespoň existenci zájmu o narušení zavedeného chápání dat.

Závěr

Epistemologii dat je třeba věnovat pozornost kvůli zásadnímu statusu připisovanému datům v současné vědě. Díky dramatickému poklesu nákladů na produkci rozsáhlých a poměrně kvalitních dat se tak data stávají ústředním zdrojem výzkumné činnosti. Pokud je datům propůjčena tato role, pak je nelze přijímat jako nezpochybnitelný koncept postavený mimo diskusi. Navzdory tomu je převládající pojetí epistemologie dat v soudobé vědě založeno na této preanalytické povaze dat. V populární pyramidě *data - informace - znalosti - moudrost* jsou data zařazena na prvním místě, jako podklad předcházející další stupně vědění. Data tak předchází leda samotná skutečnost, za jejíž přímou reprezentaci jsou pokládána.

O odhalení slabin této koncepce se zasloužilo mnoho kritiků, na jejichž díle je tento text vystavěn. Mnoho z publikací zde citovaných přivádí pozornost ke slabinám dědictví pozitivistické epistemologie. Narůstající počet autorů se začíná věnovat zpochybnění zavedené role dat ve vědě. O reformulaci epistemologie datové náročné vědy se pokouší řada výzkumníků a objevují se i první projekty zaměřené na toto téma.² O napadení dosavadního neproblematického vnímání dat se pokusil i tento text. V souladu se slovy Kenta [25, s. 220] se zde:

“představuje filosofie, že život a skutečnost jsou v podstatě amorfní, neuspořádané, protiřečící si, nekonzistentní, iracionální a neobjektivní. Věda a velká část západní filosofie nám však v minulosti vykreslovala iluzi, že s věcmi je to právě naopak.”

Kritická reflexe dominantní epistemologie dat v západní filosofii našla v nekritickém, pozitivistickém přístupu k datům řadu trhlin, v jejichž světle výklad dat jakožto nezpochybnitelného obrazu objektivně existující skutečnosti neobstojí. Jako produktivnější se proto jeví o datech přemýšlet v rámci alternativních přístupů, které nabízí například konstruktivistická epistemologie nebo rétorika. Ať už je zvolena kterákoli cesta, tak s daty ve vědě nelze zacházet jako s neproblematickým zadáním matematické úlohy, avšak je zapotřebí je podrobit otázkám.

¹Například sborník textů *‘Raw data’ is an oxymoron* z roku 2013 [18].

²Např. <http://www.datastudies.eu/>

Reference

- [1] ANDERSON, Chris. The end of theory: the data deluge makes the scientific method obsolete. *Wired* [online]. 2008-06-23 [cit. 2013-11-23]. Dostupný z WWW: http://www.wired.com/science/discoveries/magazine/16-07/pb_theory
- [2] BABBAGE, Charles. *Reflections on the decline of science in England and on some of its causes*. London: B. Fellowes; J. Booth, 1830. Dostupný také z WWW: <https://archive.org/details/reflectionsonde00mollgoog>
- [3] BACHELARD, Gaston. *The formation of the scientific mind: a contribution to a psychoanalysis of objective knowledge*. Přeložila Mary MCALLESTER JONES. Manchester: Clinamen Press, 2002. ISBN 1-903083-20-6.
- [4] BIRD, Alexander. The epistemology of science: a bird's-eye view. *Synthese*. 2010, sv. 175, č. 1 dodatek, s. 5-16. Dostupný také z WWW: http://eis.bris.ac.uk/~plajb/teaching/The_Epistemology_of_Science.pdf. DOI 10.1007/s11229-010-9740-4.
- [5] BOELLSTORFF, Tom. Making big data, in theory. *First Monday* [online]. 2013 [cit. 2013-11-28], sv. 18, č. 10. Dostupný z WWW: <http://uncommonculture.org/ojs/index.php/fm/article/view/4869/3750>
- [6] BOGEN, James; WOODWARD, James. Saving the phenomena. *The Philosophical Review*. 1988, sv. 97, č. 3, s. 303-352. Dostupný také z WWW: <http://www.pitt.edu/~rtjbog/bogen/saving.pdf>
- [7] BOX, George E. P.; DRAPER, Norman R. *Empirical model-building and response surfaces*. Hoboken (NJ): Wiley, 1987. Wiley series in probability and statistics, sv. 157. ISBN 0-471-81033-9.
- [8] BOWKER, Geoffrey C. *Memory practices in the sciences*. Cambridge (MA): MIT Press, 2005, 280 s. Inside technology. ISBN 978-0-262-52489-6.
- [9] BOYD, Danah; CRAWFORD, Kate. Critical questions for big data. *Information, Communication & Society*. 2012, sv. 15, č. 5, s. 662-679. Dostupný také z WWW: <http://dx.doi.org/10.1080/1369118X.2012.678878>. DOI 10.1080/1369118X.2012.678878.
- [10] BRUMMETT, Barry. Three meanings of epistemic rhetoric. *Speech Communication Association Convention: Seminar on Discursive Reality*. San Antonio (TX): 1979. Dostupný také z WWW: <http://ap2008.wdfiles.com/local--files/selected-research-articles/Brummett1979.doc>
- [11] CRAWFORD, Kate. The hidden biases in big data. *Harvard Business Review Blog Network* [online]. 1. dubna 2013 [cit. 2013-11-11]. Dostupný z WWW: <http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/>
- [12] DEWEY, John. *The quest for certainty: a study of the relation of knowledge and action*. New York: Minton, Balch & Company, 1929. Gifford lectures. Dostupný také z WWW: <https://archive.org/details/questforcertaint032529mbp>
- [13] DIAKOPOULOS, Nick. *The rhetoric of data* [online]. July 25, 2013 [cit. 2013-11-22]. Dostupný z WWW: <http://www.nickdiakopoulos.com/2013/07/25/the-rhetoric-of-data/>
- [14] EDWARDS, Paul N. *A vast machine: computer models, climate data, and the politics of global warming*. Cambridge (MA): MIT Press, 2010, 552 s. ISBN 978-0-262-01392-5.
- [15] EDWARDS, Paul N. [et al.] (eds.). *Knowledge infrastructures: intellectual frameworks and research challenges* [online]. Report of a workshop sponsored by the National Science Foundation and the Sloan Foundation, University of Michigan School of Information, 25-28 May 2012. May 2013 [cit. 2013-11-22]. Dostupný z WWW: <http://hdl.handle.net/2027.42/97552>

- [16] EUKLEIDÉS. Data. In SIMSON, Robert (ed.). *The elements of Euclid*. Philadelphia: De-silver, Thomas & co., 1834.
- [17] FANELLI, Daniele. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *Public Library of Science ONE* [online]. May 29, 2009 [cit. 2013-12-03], sv. 4, č. 5. Dostupný z WWW: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0005738>. DOI 10.1371/journal.pone.0005738.
- [18] GITELMAN, Lisa (ed.). *'Raw data' is an oxymoron*. Cambridge (MA): MIT Press, 2013. ISBN 978-0-262-51828-4.
- [19] GROSS, Alan G.; HARMON, Joseph E.; REIDY, Michael. *Communicating science: the scientific article from the 17th century to the present*. New York (NY): Oxford University Press, 2002. ISBN 0-19-513454-0.
- [20] HALAVAIS, Alexander. Home made big data? Challenges and opportunities for participatory social research. *First Monday* [online]. 2013 [cit. 2013-11-28], sv. 18, č. 10. Dostupný z WWW: <http://uncommonculture.org/ojs/index.php/fm/article/view/4876/3754>
- [21] HALEVY, Alon; NORVIG, Peter; PERREIRA, Fernando. The unreasonable effectiveness of data. *Intelligent Systems*. 2009, sv. 24, č. 2, s. 8–12. Dostupný také z WWW: <http://static.googleusercontent.com/media/research.google.com/en/pubs/archive/35179.pdf>. ISSN 1541-1672. DOI 10.1109/MIS.2009.36.
- [22] HENNIG, Christian. Confronting data analysis with constructivist philosophy. In *Classification, clustering, and data analysis: recent advances and applications, part II*. Berlin; Heidelberg: Springer, 2002, s. 235-243. ISBN 978-3-642-56181-8. DOI 10.1007/978-3-642-56181-8_26.
- [23] KEIL, Petr. *Data-driven science is a failure of imagination* [online]. January 2, 2013 [cit. 2013-11-22]. Dostupný z WWW: <http://www.petrkeil.com/?p=302>
- [24] KELLER, Evelyn Fox. *Reflections on gender and science*. New Haven (MA): Yale University Press, 1985. ISBN 0-300-06595-7.
- [25] KENT, William. *Data and reality*. Bloomington (IN): 1st Books Library, 2000. ISBN 1-58500-970-9.
- [26] LEONELLI, Sabina. On the locality of data and claims about phenomena. *Philosophy of Science*. 2009, sv. 76, č. 5, s. 737–749. Dostupný také z WWW: <https://ore.exeter.ac.uk/repository/handle/10871/9429>. ISSN 0031-8248.
- [27] LEONELLI, Sabina. Data interpretation in the digital age. *Perspectives on Science* [v tisku]. 2014. Dostupný také z WWW: <https://ore.exeter.ac.uk/repository/handle/10036/4484>. ISSN 1063-6145.
- [28] MAGEE, Liam. Frameworks for knowledge representation. In COPE, Bill; KALANTZIS, Mary; MAGEE, Liam (eds.). *Towards a semantic web: connecting knowledge in academic research*. Oxford: Chandos, 2011. ISBN 978-1-84334-601-2.
- [29] MARKHAM, Annette N. Undermining 'data': a critical examination of a core term in scientific inquiry. *First Monday* [online]. 2013 [cit. 2013-11-22], sv. 18, č. 10. Dostupný z WWW: <http://uncommonculture.org/ojs/index.php/fm/article/view/4868/3749>. DOI 10.5210/fm.v18i10.4868.
- [30] MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. *Big data: a revolution that will transform how we live, work, and think*. Boston (MA): Houghton Mifflin Harcourt, 2013. ISBN 978-0-544-00269-2.
- [31] Misuse of statistics. *Wikipedia* [online]. Naposledy změněno 6. prosince 2013 [cit. 2013-12-08]. Dostupný z WWW: http://en.wikipedia.org/wiki/Misuse_of_statistics

- [32] NELSON, Michael L. Data-driven science: a new paradigm? *EDUCAUSE Review* [online]. July/August 2009 [cit. 2013-11-22], sv. 44, č. 4, s. 6–7. Dostupný z WWW: <http://www.educause.edu/ero/article/data-driven-science-new-paradigm>
- [33] NIELSEN, Michael. *Reinventing discovery: the new era of networked science*. New Jersey: Princeton University Press, 2011, 273 s. ISBN 978-0-691-14890-8.
- [34] NUNBERG, Geoffrey. Farewell to the Information age. In NUNBERG, Geoffrey (ed.). *The future of the book*. Berkeley (CA): University of California Press, 1996. ISBN 0-520-20451-4.
- [35] PELED, Alon. *The politics of big data: a three-level analysis* [online]. 2013 [cit. 2013-12-06]. Dostupný z WWW: <http://ssrn.com/abstract=2315891>
- [36] PIAGET, Jean. The psychogenesis of knowledge and its epistemological significance. In PIATTELLI-PALMARINI, Massimo (ed.). *Language and learning: the debate between Jean Piaget and Noam Chomsky*. Cambridge (MA): Harvard University Press, 1980. ISBN 0-674-50940-4.
- [37] POOVEY, Mary. *A history of the modern fact: problems of knowledge in the sciences of wealth and society*. Vyd. 1. Chicago: University of Chicago Press, 1998, 436 s. ISBN 0-226-67526-2.
- [38] PORTER, Theodore M. *Trust in numbers: the pursuit of objectivity in science and public life*. Princeton (NJ): Princeton University Press, 1995. ISBN 0-691-03776-0.
- [39] SAPIR, Edward. *The collected works of Edward Sapir. VIII, Takelma texts and grammar*. Berlin; New York: Mouton de Gruyter, 1990. Dostupný také z WWW: <https://archive.org/details/collectedworksof01sapi>
- [40] SCHRON, Max. *Data's missing ingredient? Rhetoric* [online]. 11. 4. 2013 [cit. 2013-12-09]. Dostupný z WWW: <http://strata.oreilly.com/2013/04/datas-missing-ingredient-rhetoric.html>
- [41] SCOTT, Robert L. On viewing rhetoric as epistemic. *Central States Speech Journal*. 1967, sv. 18, č. 1, s. 9-17. DOI 10.1080/10510976709362856.
- [42] SPRAT, Thomas. *The history of the Royal-Society of London, for the improving of natural knowledge*. [T.R.]: London, 1667, 438 s. Dostupný také z WWW: <https://archive.org/details/details/historyroyalsoc00martgoog>
- [43] TOULMIN, Stephen E. *The uses of argument*. Cambridge: Cambridge University Press, 2003. ISBN 978-0-511-07117-1.
- [44] VON GLASERSFELD, Ernst. An introduction to radical constructivism. In WATZLA-WICK, Paul (ed.). *The invented reality*. New York: Norton, 1984, s. 17-40.

E-LOGOS

ELECTRONIC JOURNAL FOR PHILOSOPHY

Ročník/Year: 2013 (vychází průběžně/ published continuously)

Místo vydání/Place of edition: Praha

ISSN 1211-0442

Vydává/Publisher:

Vysoká škola ekonomická v Praze / University of Economics, Prague

nám. W. Churchilla 4

Czech Republic

130 67 Praha 3

IČ: 61384399

Web: <http://e-logos.vse.cz>

Redakce a technické informace/Editorial staff and technical information:

Miroslav Vacura

vacuram@vse.cz

Redakční rada/Board of editors:

Ladislav Benyovszky (FHS UK Praha, Czech Republic)

Ivan Blecha (FF UP Olomouc, Czech Republic)

Martin Hemelík (VŠP Jihlava, Czech Republic)

Angelo Marocco (Pontifical Athenaeum Regina Apostolorum, Rome, Italy)

Jozef Kelemen (FPF SU Opava, Czech Republic)

Daniel Kroupa (ZU Plzeň, Czech Republic)

Vladimír Kvasnička (FIIT STU Bratislava, Slovak Republic)

Jaroslav Novotný (FHS UK Praha, Czech Republic)

Jakub Novotný (VŠP Jihlava, Czech Republic)

Ján Pavlík (editor-in-chief) (VŠE Praha, Czech Republic)

Karel Pstružina (VŠE Praha, Czech Republic)

Miroslav Vacura (executive editor) (VŠE Praha, Czech Republic)