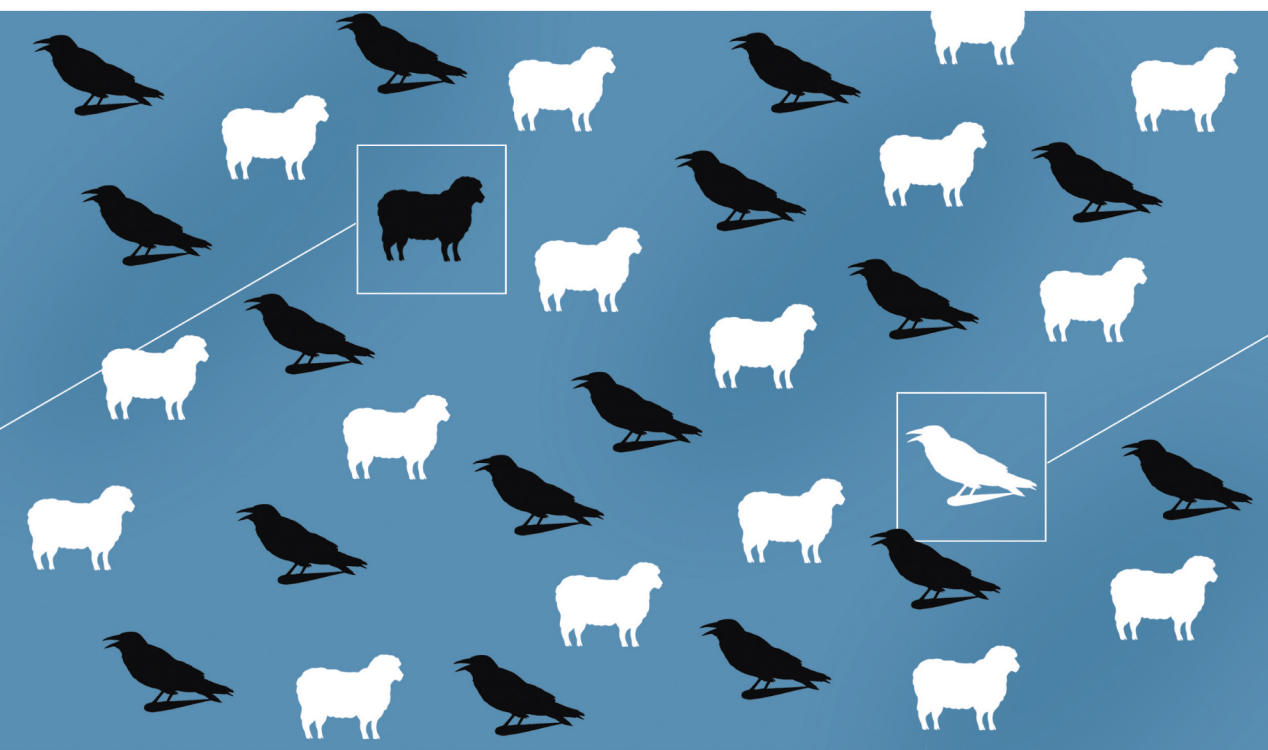


Zach Gemignani, Chris Gemignani,
Richard Galentino, Patrick Schuermann

Efektivní analýza a využití dat



Příprava dat pro různé cílové skupiny
Účinné metody prezentace
Grafy a vizualizace dat
Tipy uznávaných autorů

computer
press

WILEY

**Zach Gemignani
Chris Gemignani
Richard Galentino
Patrick Schuermann**

Efektivní analýza a využití dat

**Computer Press
Brno
2015**

Efektivní analýza a využití dat

Zach Gemignani, Chris Gemignani, Richard Galentino, Patrick Schuermann

Překlad: Jiří Huf

Obálka: Martin Sodomka

Odpovědný redaktor: Libor Pácl

Technický redaktor: Jiří Matoušek

Authorized translation from the English language edition Data Fluency: Empowering Your Organization with Effective Data Communication.

Copyright © 2014 by John Wiley & Sons, Inc., Indianapolis, Indiana

All Rights Reserved. This translation published under license with the original publisher John Wiley & Sons, Inc.

Translation © Jiří Huf, 2015

Objednávky knih:

<http://knihy.cpress.cz>

www.albatrosmedia.cz

eshop@albatrosmedia.cz

bezplatná linka 800 555 513

ISBN 978-80-251-4571-5

Vydalo nakladatelství Computer Press v Brně roku 2015 ve společnosti Albatros Media a. s. se sídlem Na Pankráci 30, Praha 4. Číslo publikace 19 283.

© Albatros Media a. s. Všechna práva vyhrazena. Žádná část této publikace nesmí být kopírována a rozmnožována za účelem rozšiřování v jakékoli formě či jakýmkoli způsobem bez písemného souhlasu vydavatele.

1. vydání

 **ALBATROS MEDIA** a.s.

Obsah

O autorech	11
Poděkování	13
Předmluva	15
Úvod	17
Data jsou novým obchodním jazykem	18
Datová komunikace není technologický, ale společenský problém	19
Spolupráce	20
Od dat k činům	20
Vizualizace je pouze částí řešení	21
Pro koho je tato kniha?	22
Poznámky	23

KAPITOLA 1

Problém poslední míle	25
Informační doba: potřeba přenášet data plynule	26
Plynulost dat: Využijte potenciál svých firemních dat	28
Veledata a datová přeměna	28
Náš rámec pro plynulý přenos dat	30
Případové studie: Okno do rámce plynulosti předáváníí dat	31
Poznámky	39

KAPITOLA 2

Rámec plynulého přenosu dat	41
Rámec plynulého přenosu dat	43
Jednotlivci a společnost	44
Používání versus prezentování dat	44
První element: Datově gramotní konzumenti	45
Druhý element: Tvůrci schopní přenášet data plynule	46
Třetí element: Kultura plynulého přenosu dat	47
Čtvrtý element: Ekosystém datového produktu	48

Pojící nit	49
Kam za podrobnějšími informacemi	49
Výhody plynulého přenosu dat ve společnosti	50
Jak s rámcem pracovat	50
Jak společnosti bojují	51
Shrnutí	52
Poznámky	52

KAPITOLA 3

Jak společnosti bojují s plynulým přenosem dat	53
Překážky na cestě k plynulému přenosu dat	54
Množící se reporty	55
Balkanizovaná data	56
Datová elita	57
Dokonalý model	58
Hledání porozumění	59
Péče o data	60
Lpění na konkrétním ukazateli	61
Hledání rovnováhy	63
Poznámky	64

KAPITOLA 4

Konzumentův průvodce po datech	65
Datové výstupy	67
Běžné datové výstupy	68
Překážky v práci s datovými výstupy	72
Když nevíte, kde začít a na co se zaměřit	74
Nekonzistence	75
Učíme se jazyku dat	76
Atomická data	77
Souhrnná data	77
Průzkum datových tabulek	78
Grafy a vizualizace	80
Srozumitelnost	81
Rozklad datových výstupů	82

Co s tím můžete dělat?	88
Shrnutí	91
Poznámky	92

KAPITOLA 5

Autoři dat: Zkušení návrháři datových prezentací	93
Vzácné dovednosti	94
Co se naučíte	95
Řízený rozhovor	96
Hledání účelu a významu	97
Výběr údajů	101
Definování smysluplných a použitelných metrik	104
Vytváření struktury a tok datových výstupů	106
Tvorba přitažlivých a srozumitelných datových výstupů	114
Navazování dialogu datovými výstupy	122
Pravidla designu	123
Vizualizace	123
Pravidla designování datových výstupů	129
Ať žijí autoři datových výstupů	133
Poznámky	134

KAPITOLA 6

Kultura plynulého přenosu dat	135
Vedení, kultura a priority v komunikaci	136
Stanovte očekávání	137
Definujte konkrétní, měřitelné ukazatele	137
Chvalte efektivní výstupy a práci s daty	138
Definování klíčových metrik	140
Jak vypadá dobrá metrika?	140
Zlepšování společnosti pomocí metrik	143
Ukažte zaměstnancům, že je jejich přínos stěžejní	143
Společné porozumění	146
Společný slovník a terminologie vztahující se k datům specifickým pro danou společnost	147
Porozumění důvěryhodným a spolehlivým zdrojům dat a jejich ocenění	148
Manipulace s daty a jejich modelování musí být transparentní	150

Definujte společné klíčové metriky	151
Porozumění účelu datových výstupů	152
Každodenní aktivity	153
Datoví konzumenti	153
Datové výstupy	154
Využití dat	157
Vývoj kultur s plynulým přenosem dat	158
Poznámky	159

KAPITOLA 7

Ekosystém datových výstupů	161
Datové výstupy předávající informace	162
Nezbytné podmínky	163
Učíme se v App Storu	165
Poptávka	166
Hierarchická mapa potřeb	166
Potřeby řadových zaměstnanců	167
Kde začít	168
Design	168
Cíl	169
Začněte s průvodcem stylů	170
Rozvoj	170
Jen špatný řemeslník svaluje vinu na nástroje	172
Objevování	173
Cíl	173
Kde začít: centralizovaný inventář datových výstupů	174
Diskuze	174
Cíl	174
Kde začít: založte místo pro ukládání poznatků	175
Filtrování	176
Učíme se z Wikipedie	176
Cíl	176
Bez čeho se objedete?	177
„Nalezněte si k nim vztah“	177
Poznámky	178

KAPITOLA 8

Cesta k plynulému přenosu dat	179
Proč plynulý přenos dat?	181
Datoví konzumenti: Výchova sofistikovaného publika	183
Tvůrci datových výstupů: dovednosti potřebné k efektivní datové komunikaci	183
Kultura plynulého přenosu dat: jednotné porozumění datům	184
Ekosystém datového výstupu: nástroje a procesy zprostředkující výměnu informací	185
Začátek cesty	185
Poznámky	186

VÝŇATEK

Inventář plynulého přenosu dat	187
Dotazníkové otázky inventáře plynulého přenosu dat	188
První komponenta: gramotnost konzumentů dat	189
Druhá komponenta: dovednosti tvůrců datových výstupů	192
Třetí komponenta: kultura plynulého přenosu dat	195
Čtvrtá komponenta: ekosystém datových výstupů	197
Shrnutí Inventáře	199
Jak hodnotit (společnosti)	199
Typy otázek a bodové hodnoty	199
Hodnocení společnosti	200
První komponenta: gramotnost konzumentů dat	200
Druhá komponenta: tvůrci datových výstupů	200
Třetí komponenta: kultura plynulého přenosu dat	201
Čtvrtá komponenta: ekosystém datových výstupů	202
Shrnutí hodnocení společnosti	203
Jak hodnotit (jednotlivce)	203
Co můžete coby jednotlivce měřit?	203
První komponenta: konzument dat (pouze jednotlivce)	204
Druhá komponenta: autoři datových výstupů (pouze jednotlivce)	204
Shrnutí hodnocení jednotlivce	205
Materiály pro Inventář	205
Doprovodný e-mail dotazníku	205
Kvíz datové gramotnosti	206

PŘÍLOHA A

Design datových výstupů	211
Články	211
Kontrolní seznam pro tvorbu datových výstupů	212
Myslete jako návrhář	214
Navržen k použití	216
Opouštíme pravidlo jednostránkových přehledů	217
Kontrolní seznam upozornění na přehledu	219
Kontext: Uživatelé musí chápat, co je upozornění a jak zapadá do celku	220
Přesvědčivost: systém upozornění nesmí vyvolávat zbytečnou paniku, ale musí podávat snadno srozumitelné informace, na jejichž základě lze jednat	220
Komunikace: upozornění musí poutat pozornost a informovat efektivně	220
Řízení: pokročilé výstražné systémy by měly uživateli umožnit si upozornění přizpůsobit a spravovat je	221
8 znaků kvalitních přehledů zobrazujících data v reálném čase	221
Poznámky	223

PŘÍLOHA B

Průvodce správným stylem	225
První vzorek: fonty	226
Druhý vzorek: barvy	227
Třetí vzorek: formátování dat a čísel	228
Čtvrtý vzorek: pruhové grafy	229
Pátý vzorek: grafy znázorňující trendy	230
Šestý vzorek: tabulky	231
Rejstřík	233

Knihu věnujeme svým rodičům, dělili se s námi o zálibu v umění a potěšení z výuky.
My se totéž snažíme předat těm, kteří pracují s daty.

O autorech

Tato kniha je kolektivní prací, která staví na letech zkušeností, jež jsme získali, když jsme pomáhali firmám nakládat lépe s daty. Zach Gemignani měl na starosti psaní textu a definoval rámec plynulého přenosu dat, který je základem této knihy. Chris Gemignani dal podnět mnohým vizualizacím dat, designu i našim společným myšlenkám. Richard Galentino přispěl svými zkušenostmi v oblasti zdravotní péče, vzdělávání a neziskových organizací, stojí za inventářem plynulého přenosu dat a měl na starosti koordinaci s redaktory. Patrik Schuermann spolupracoval s naším pomocným průzkumníkem Timem a spolu dodali materiál týkající se vývoje společností. Díky nim se z naší knihy stal nástroj pro vedoucí pracovníky, kteří chtějí svou společnost změnit.

Zach Gemignani je zakladatelem společnosti Juice Analytics. Této společnosti pomohl vystavět jméno na poli designu zajímavých údajů a jedinečných řešení pro vizualizaci dat. Coby výkonný ředitel zodpovídá za strategii směřování společnosti, přináší nové myšlenky a rozvíjí obchodní příležitosti společnosti. Dříve se věnoval datovým sestavám a analýze ve společnosti AOL a působil jako konzultant u společností Diamond Technology Partners a Booz Allen, u nichž byl znám jako tvůrce úžasných prezentací. Promoval na Haverfordu coby bakalář ekonomie. Titul MBA si vysloužil na The Darden School virginské univerzity. Nyní žije v Nashvillu v Tennessee s manželkou a třemi dětmi.

Chris Gemignani je spoluzakladatelem společnosti Juice Analytics a současně vizionářem společnosti. Chris si vysloužil uznání v oblasti kreditních karet. Zodpovídal za modelování rizik a analýzu chování držitelů platebních karet. Ke své zkušenosti s touto analýzou přidává schopnost předat své poznatky dále, ale také kritický pohled na uživatelské rozhraní a design interaktivních prostředí. Bakalářský titul získal na Williams College v oboru počítačové vědy a ekonomie. Magisterský titul pak získal na Washingtonské univerzitě v St. Louis v oboru ekonomie.

Dr. Richard Galentino pracuje jako výkonný ředitel společnosti Stratable, Inc., což je společnost zabývající se strategickým plánováním a rozvojem společnosti. Dříve vedl mezinárodní tým, který vysílal stovky lékařů, sester a přidruženého zdravotnického personálu do více než 27 zemí. Richard má za sebou harvardskou univerzitu, kde se účastnil programu Education Policy Fellowship. Zde získal magisterský titul na poli vzdělávání, správy, plánování a sociální politiky. Na School of Foreign Service Georgetownské univerzity pak získal bakalářský titul v oboru ekonomie. Doktorát získal v oblasti vedení vzdělávání a veřejné správy na Vanderbiltově univerzitě. Richard žije společně se svou rodinou v Nashvillu v Tennessee.

Dr. Patrick Schuermann je profesorem pro výzkum na Peabody College of Education Vanderbiltovy univerzity. Dříve působil jako ředitel pro technickou pomoc a řízení národního centra pro reformu odměňování učitelů zřízeného ministerstvem školství Spojených států amerických, ale také jako hlavní výzkumník v řadě projektů věnovaných řízení školy a vzdě-

lávací technologii. Momentálně působí jako ředitel nezávislého magisterského programu pro řízení škol a předseda profesionálních institutů Peabody College of Education. Sídlí v Nashvillu společně se svou talentovanou manželkou (písničkářkou) a dvěma psy.

Poděkování

Před deseti lety jsme se společně s Chrisem rozhodli, že si založíme firmu a vyplníme mezeru mezi daty a lidmi, kteří s daty pracují. Tato kniha vychází z četných zkušeností, které jsme o prezentaci, vizualizaci a sdílení dat posbírali při diskuzích s kolegy a klienty.

Jsem vděčný za nasazení, se kterým se do práce pustili všichni mí spolupracovníci ve společnosti Juice Analytics. Zvláště děkuji Kenovi Hilburnovi, Jamesovi Lytleovi a Michelovi Guilletovi, kteří mi pomohli při psaní přemýšlet. Do práce se hned chodí s radostí, když můžete pracovat s tak nadanými lidmi, jako jsou Djam Saidmuradov, Meghna Kukrejaová, Lindsay Concharová, Tim O'Guin a Jennie Gemignaniová.

Rád bych poděkoval Nathanovi Yauovi, který nás v nakladatelství Wiley přivítal s otevřenou náručí. Jeho neutuchající nadšení pro učení, sdílení a diskuzi o všech aspektech vizualizace dat je nejen úchvatné, ale i podnětné. Adaobi Obi Tultonová nás coby projektová redaktorka trpělivě provedla všemi fázemi publikace knihy.

Rovněž jsem vděčný za pozornou pomoc Tima Drakea, který se podílel na textu knihy i na výzkumu. Tim je doktorand a studuje vedení a řízení vzdělávání na Vanderbiltově univerzitě. Píše, zkoumá a vyučuje v oblastech designu a metod kvantitativního výzkumu, rozhodování na základě dat a vedení a řízení vzdělávání na základním a středním stupni.

A konečně poslední srdečné poděkování patří mé ženě Andree a dětem – Owenovi, Mayae a Lile –, kteří mi byli podporou a trpělivě snášeli, že si nabírám další práci.

Zach

Předmluva

Od posledního statistického semináře, na který jsem na výšce chodil, je to již více než deset let. Na rozdíl od mnohých mám na začátky se statistikou hezké vzpomínky. Učil mě nadšený profesor, který neustále chodil po schodech v aule nahoru a dolů. Základní principy distribuce a ověřování hypotéz vás tak snadno nenadchnou, ale on to dokázal. Čím dál více mě bavilo pracovat s daty. Chtěl jsem jim porozumět, což mě nakonec vedlo k dalším rokům studia. Na škole jsem ještě neměl vůbec ponětí, že se statistika – nebo obecněji práce s daty – stane natolik populární. Prostě jsem si jen rád hrál s daty. A dnes máme k dispozici opravdu hodně dat.

Každodenně jsem četl nebo slyšel o společnostech, které nějakým způsobem pracovaly s daty. Data lze použít celou řadou způsobů: mohou pomoci obchodu, poskytnout lepší služby zákazníkům, usnadnit život dalším lidem, ale také pomáhají při znázorňování složitých procesů. Mají své kouzlo. Lidé chtějí mít užitek ze všech dat, která sesbírají.

Má to ale háček, a to dost velký. Nestačí jen vzít nějaký tok dat a zapojit ho do nejdražšího softwaru, jaký seženete, a ihned čerpat výsledky – a to bez ohledu na to, jestli pracujete jako jednotlivec či jako společnost. Takhle jednoduché to rozhodně není. A každý, kdo vám bude tvrdit opak, buď vůbec neví, o čem mluví, nebo se vám snaží něco prodat.

Coby člověk, který se zaměřuje na vizualizaci dat, bych rád sestavil nějaký řídicí panel nebo nástroj, jenž by umožnil lidem porozumět datům ihned. Aniž by byli znalci. Jenže abyste dokázali používat jakýkoliv nástroj účelně, musíte se s ním nejprve naučit pracovat. Musíte mít představu o tom, o čem data vypovídají, jak je analyzovat a vykládat.

Když se začnete soustředit na to, jak se může při práci s daty zlepšit celá společnost, narazíte na další překážky. Vedení nese zodpovědnost za něco jiného než řadoví pracovníci, ale aby spolu mohli všichni pracovat efektivně, musí vycházet ze společných základů.

Zach a Chris Gemignaniovi, spoluzakladatelé společnosti Juice Analytics, pomáhají řešit takové situace dennodenně a učí své žáky tomu, co nazývají plynulým přenosem dat (Data Fluency). Oba bratři spolu se svými spolupracovníky pracovali jako konzultanti již dávno předtím, než se objevila tzv. veledata (big data). Dříve, než hlavní ekonom společnosti Google Hal Varian řekl, že práce statistika je sexy, a dříve, než jsem založil společnost FlowingData. Jejich zkušenosti jsou v článcích na Internetu i v této knize dobře vidět. Dávají praktické, ovšem dostatečně obecné rady, které pomáhají nasazovat nabízená řešení v konkrétní situaci.

Když jsem před lety poprvé hledal heslo „vizualizace dat“, společnost Juice Analytics byla mezi prvními, na něž jsem narazil, a dodnes jsem jejich odběratelem. Měl jsem proto radost, když se Zach s Chrisem nabídl, že tuto knihu napíší. Tato kniha však není knihou o vizualizaci. Rozhodně se tomuto tématu věnuje, nicméně podává širší pohled na věc.

Pokud ve své firmě používáte vizualizaci – reporty, prezentace a zobrazení dat –, vadí vám, když je nikdo nesleduje ani z nich nemá užitek? Když skončí v koši anebo když jejich výkladu

naslouchají posluchači jen mimoděk? Můžete připravit nejlepší graf na světě, ale jde o to, aby vám posluchači věnovali pozornost. Je důležité nechat data vyniknout, aby se všichni, vedením počínaje, mohli rozhodovat na nějakém faktickém základě.

Je ale pravdou, že neexistuje žádný zázračný lék či nástroj, který by vyřešil všechny problémy pojící se k datům. Zkušenosti bratří Gemignaniů nám podávají realistický pohled. Oba bratři chápou, že některé části celku se mohou pohybovat pomaleji a nemusí být přístupné změně. Ve společnosti FlowingData sice pracuji sám, ale v roli konzultanta jsem již dobře pochopil, že byrokracie má své stinné stránky. Klíčem je začít tam, kde je změna možná. Tato kniha je skvělým návodem, který vám umožní pochopit, jak to udělat.

A když tu tak sedím a představím si, jak budou data vypadat za deset let, vidím jich více a složitěji poskládaných. V současnosti sbíráme data výrazně rychleji, než jim dokážeme porozumět. Jsme však svědky rostoucí touhy lidí datům porozumět. Takže čím více lidí, kteří se chtějí naučit jazyk dat, budeme mít dnes, tím lépe pro budoucnost.

Nathan Yau

Úvod

Jak lze změnit lidi?

S bratrem jsme se tísnilí ve sklepech a dokončovali jsme analýzu. Slunce už zapadlo a druhý den jsme měli prezentovat. Celý poslední měsíc jsme strávili sběrem dat popisujících míru setrvání studentů v online kurzech. Chtěli jsme vědět, proč studenti z kurzů odchází a kteří studenti mají tendenci vytrvat.

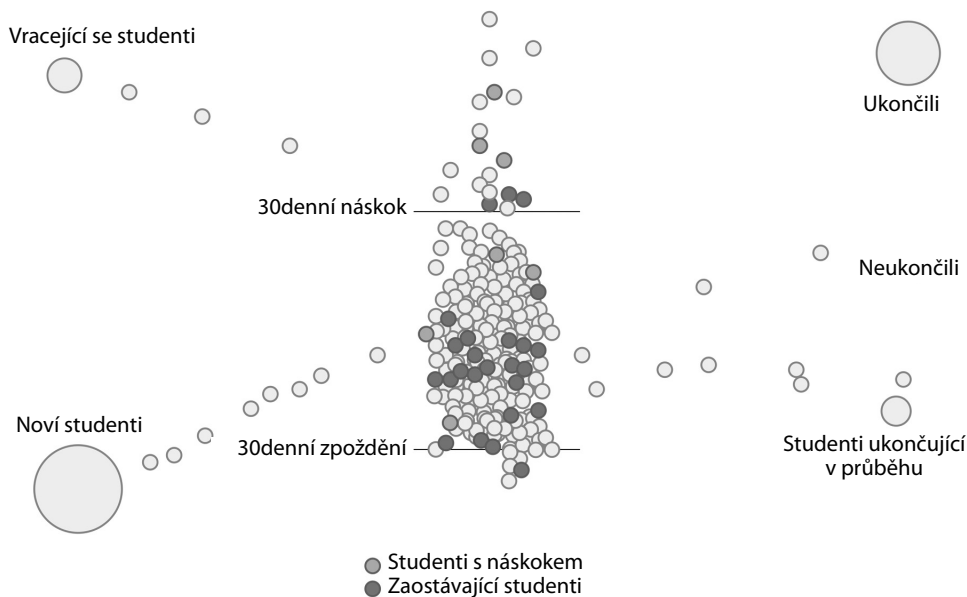
Připravili jsme si snímky, které jsme chtěli předvést vedení. Naše prezentace předváděla model, který populaci studentů rozděloval a nabízel doporučení. Měl jsme však pocit, že nám něco chybí.

Jak posluchače zaujmout, aby měli o prezentaci zájem?

Za našimi čísly stáli konkrétní studenti, kteří si zvolili online školu, vzali si studentské půjčky, aby mohli zaplatit školné, strávili hodiny s domácími úkoly, hovořili se svými vyučujícími a snažili se vše včas zvládnout. Naše analýza však jejich situaci, úspěchy i snahu jen zploštila. Jak přenést prezentaci zpět do reality?

Rafičky na hodinách se již blížily půlnoci a my dostali nápad. Připravíme animovaný film. Zachytíme v něm, jak se jednotliví studenti dostali do škol, jak dělali domácí úkoly a jak se nakonec rozhodli odejít či zůstat.

Film jsme spíchli horkou jehlou. Dali jsme dohromady obrázky s jednotlivými studenty, které jsme v každém školním roce spojili dohromady (obrázek 1). Studenty jsme přesně rozmístili a pohybovali jsme s nimi jako se snímkanými postavičkami. Chris nakonec napsal skript, který zachytil stovky snímků jednotlivých dní. Ty pak spojil do animace. Coby unaveným nadšencům do dat nám pár kouzelných minut filmu připadalo jako Spielbergův majstrštyk.



Obrázek 1: Snímek z našeho filmu o studentech v online kurzech

Studenti pochodovali po obrazovce na cestě ke zdárnému ukončení či k lítostnému odchodu z kurzu. Data dostala nový význam. A nastartovala diskuzi s naším klientem.

Kreativní provedení vybudilo vášně. Juice Analytics jsme založili o několik měsíců dříve. Věděli jsme, že chceme pomoci firmám pochopit vlastní data.

Díky tomu večeru jsme se dokázali zaměřit nejen na čísla, ale také na to, jak jsou prezentována. Uvědomili jsme si, že potřebujeme najít lepší a kreativnější způsob, jak lidem pomoci porozumět datům. Dokážeme postavit most mezi analytiky a mezi lidmi, kteří se rozhodují na základě jejich práce?

Za tímto cílem jsme šli skoro deset let. Juice Analytics mezitím spolupracovala se stovkami společností – od startupů, které se snažily předkládat data svým zákazníkům, až po celosvětové značky, které se snažily najít způsob, jak data lépe předkládat svým vedoucím pracovníkům. Navrhovali jsme zajímavé interaktivní řídicí panely, reporty a analytické nástroje – a vždy jsme se přitom snažili pomáhat skutečným lidem porozumět datům a pracovat s nimi.

A za tu dobu jsme získali pár podstatných poznatků.

Data jsou novým obchodním jazykem

Data jsou médiem, které slouží ke komunikaci a přesvědčování. V posledních letech jsme se dobře seznámili s jejich hodnotou a dále ji posilujeme. Na mediálních stránkách, například FiveThirtyEight (ESPN) a Upshot (The New York Times), se zakládají veřejná diskuzní fóra,

v nichž se datová analýza a vizualizace snoubí s žurnalistikou. Tyto stránky jsou veřejným vyjádřením trendu, jenž prosakuje mnoha moderními organizacemi.

Ne každému se však pomocí dat komunikuje dobře. Publikum, pro něž navrhujeme – systémoví správci, právníci, marketéři –, si často neumí z dat vzít ty správné informace, a to ani po kouscích. Data nyní potřebují otevřít diskuzi. Mluvčí se musí vyjadřovat jasně a přesně, zatímco publikum musí být schopným příjemcem. Přednášející musí oslovit publikum, které před nimi právě sedí, jen tak bude posluchač schopný dobře přijmout i obsah prezentace.

Jak ale položit nějaké společné základy efektivní komunikace? Můžete začít s výukou základních pravidel vizualizace dat: metriky, dimenze, distribuce, vztahy, okraje a změny. K dobrým metodám prezentace dat se dostanete tak, že si do začátku vyberete vhodný graf, který znázorňuje důležité prvky. To, jak plynule hovořit jazykem dat, se můžete učit od zkušených mluvčích.

Data se více než kdy dříve budou stávat stále častější součástí komunikace. Je třeba se postarat o to, aby se diskuze mohli účastnit všichni ve firmě.

Datová komunikace není technologický, ale společenský problém

Dlouhé roky se mnohé firmy snažily získat obrovské objemy dat a investovaly do větších databází a na funkce bohatých platform. Zaměření na rozsah je však odvádělo od skutečného přínosu – užitečné perspektivy v rukou těch, kteří mají moc něco ovlivnit. První generace analytických nástrojů se zaměřovala na komplexní, na funkce bohatá řešení, která byla navržena pro IT specialisty. Sesbírání kvanta dat však takto zůstávají mimo dosah těch, kteří je skutečně dokáží použít.

Užitnost dat je problém, který musí vždy nakonec vyřešit konkrétní lidé – lidé, kteří rozumí konkrétnímu kontextu, lidé, kteří se rozhodují v první linii, ale také lidé, kteří dobře rozumí problémům, do nichž dokáží data vnést světlo. Data jsou užitečná tehdy, když s jejich pomocí dokáže člověk vyprávět příběh, sestavit uchvacující vizualizaci a dobře promyšlenou analýzu. Těchto lidí je však nedostatek.

Abychom odhalili skutečný přínos dat, budeme potřebovat více než jednotlivce. Je třeba zapojit interakci mezi lidmi, kteří s daty komunikují, diskutovat nad jejich významem a promyslet si další postup. Ve firmě je zapotřebí datové kultury, která se soustředí na rozhodování na základě informací.

Data jako taková jsou jen ledovým, nepřívětivým médiem. Je třeba je zlidštit a ušít lidem na míru. Je třeba je jednoznačně navázat na odpovídající problematiku, publikum pak o ně začne jevit zájem. Data se mají prezentovat intuitivně, vizuálně a jednoduše. A stejně jako u každého jiného jazyka bychom se i u dat měli snažit předat význam.

Spolupráce

Mnohé z diskuzí o datech se odehrávají na dělicí čáře mezi těmi, kdo o datech mají značné znalosti, a těmi, kteří se při práci moc často do dat neponořují. Mezi těmito skupinami lidí najdeme jazykové překážky, očekávání, ale i mylné předpoklady.

Na jedné straně si představme datového analytika, který sestavil komplexní tabulku, která slouží k popisu nedostatků v postupech anebo popisuje kompetence marketinkového kanálu. Aby tohle daný člověk dokázal, musí se seznámit se záludnostmi rozhraní API, postupy sběru a definování dat, musí se seznámit se s tím, kde lze data sbírat a jak je spojovat s ostatními daty. Pro takového člověka jsou data baterkou, která vrhá světlo na jinak chaotický svět.

Analytikova šéfová však přichází ze zcela jiného světa. Je sice stejně chytrá a stejné úsilí věnovala tomu, aby dokázala zlepšit vyhlídky společnosti, ale nemá dostatek prostoru ani času, aby se zabývala podrobnou tabulkou či algoritmy. Daleko více než tabulka s čísly ji osloví poutavý příběh.

Pokud tyto skupiny lidí dokáží najít způsob, jak spolupracovat, společnost tím jen získá. Analytikova práce se tak může dostat na světlo a vést k rozumnějším rozhodnutím. Vedoucí se pak může soustředit na aktuální problémy, u nichž může mít lepší analýza vliv na rozhodování.

Naším cílem je otevřít konstruktivní debatu mezi těmi, kteří datům dobře rozumí, a těmi, kteří se jazyku dat doposud učí. Pokud se nám to podaří, dokážeme propojit lidi, kteří dokáží klást ty nejlepší otázky, s těmi, kdo na ně dokáží odpovědět.

Od dat k činům

Přednášel jsem právě o nevyužitě hodnotě dat, když jsem si všiml, že v zadní části místnosti vystřelila nahoru ruka. Byl to podnikatel, kterého jsem znal již roky. Podnikatel ostatním řekl: „Data nemají hodnotu. Ve skutečnosti jsou jen drahá. Vemte si všechny ty peníze, které utratíte, abyste je sesbírali, uložili, roztřídili, peníze na software. Nemá naopak cenu to, že nám data pomáhají pochopit situaci?“ Velmi hezky popsal daný rozdíl. Ale možná bychom se měli pustit dále. Pravou hodnotu dat zjistíte až teprve v okamžiku, kdy se na jejich základě rozhodnete.

Firmy specializující se na data – velké společnosti v oblasti business intelligence (analýza dat), např. IBM, MicroStrategy a SAP – mají spousty důvodů, aby vás přesvědčovaly, že shromažďování dat má pro vás význam. Větší objemy dat si žádají výkonnější nástroje a komplexnější balíky funkcí. Čím více dat, tím lépe.

Nevěřte tomu.

Klíčovým pravidlem produktového designu, alespoň podle Joshuy Portera, vedoucího oddělení User Experience ve společnosti HubSpot, je: „Nejdůležitější je užitnost.“ A dále pokračuje: „Pokud váš produkt není užitečný, pokud pro něj lidé nedokáží najít použití, došlo k chybě při designu. Výrobek musí lidem pomoci dělat něco hodnotného.“¹

Data jsou užitečná v okamžiku, kdy někomu pomohou v práci vykonat něco užitečného. Když informace předáte tak, aby jim příjemci rozuměli, vyvolají diskuzi. Lidé, kteří datům porozumí, se dokáží bavit mnohem produktivněji. A z takové diskuze pak vychází lepší rozhodnutí. Když nepodniknete tuto cestu od pochopení k rozhodování, jsou data jen vzduch.

Vizualizace je pouze jednou částí řešení

Lidský mozek má neskutečnou schopnost absorbovat a zpracovávat vizuální vjemy. Nemusíme si kupříkladu zapamatovat telefonní číslo či jméno člověka, se kterým jste se právě seznámili, ale po vizuální stránce zahanbíme i počítače. Všechny ty otravné CAPTCHA systémy, které ověřují, zda nejste pouze spamující robot, jsou důkazem, že naše mysl dokáže snadno zpracovat informace a nacházet v nich vzory. To jsou základy, na nichž je vystavěna vizualizace dat.

V posledním desetiletí jsme byli přímými svědky obrovského rozkvětu vizualizace dat. V roce 2005 jsme začali o datové vizualizaci blogovat. Rodící se komunita kolem datové vizualizace tehdy tvořila jen hrstka nadšenců a akademiků. Dnes jsou již zavedeny desítky konferencí, akademických programů a tisíce designérů, kteří vám rádi zvizualizují data do infografiky.

V celém tom ruchu se však ztrácí jedna prostá skutečnost: datová vizualizace je pouze prostředkem vedoucím k cíli. Oním cílem je efektivní předávání nápadů a vhledů, a to pomocí utváření a znázorňování čísel způsobem, kterému porozumí každý. Tento prostředek dokáže – a měl by – přesahovat pouhé grafy, infografiku a složité interaktivní vizualizace.

Jak se dostaneme za práh infografiky? Pomoci publiku pracovat s daty můžeme pouze za předpokladu, že si informacemi proktestíme logickou cestu, která má jasný začátek a konec a současně vede k rozhodnutím. Předat informace správně znamená vysvětlit význam ukazatelů a včas vložit kontextuální informace. Prezentovat data neznámá pouze vizualizovat je. Jde také o to, jak s nimi dokáže uživatel pracovat, jak je může prozkoumávat a využít jich na základě toho, co již zná.

Jsou-li data novým komunikačním médiem, měli bychom se něčemu přiučit z četných forem komunikace, které jim předcházely – například z tisku, fotografie a filmu. Filmový režisér dělá mnohem více, než že jen skládá dohromady řetězec obrázků. Navazuje spojení s divákem, a to prostřednictvím umělecky pospojované hudby, zvukových efektů, střihu a kinematografie.

Role osoby, která předává data, je podobně složitá. Jejím cílem by mělo být podat „zkušenost s informacemi“, která redefinuje to, jak posluchač uvažuje o dané problematice, a pomůže mu se lépe rozhodovat.

To, zda stojí data za lepšími rozhodnutími, často závisí na dynamice společnosti. Kultura společnosti má hluboký vliv na to, zda „datové produkty“ vedou k lepším rozhodnutím. Procesy musí s efektivním předáváním dat počítat. V dynamických společnostech je efektivní předávání dat motorem, který zlepšuje procesy, systémy i rozhodnutí.

Právě o tyto poznatky, ale i mnohé další, se s vámi podělíme v této knize. Máme velká očekávání. Rádi bychom načrtli mapu, podle níž by mohly malé i velké společnosti zlepšovat práci s daty.

Pro koho je tato kniha?

Cesta k plynulému nakládání s daty je důležitá pro každou firmu, která se chce na základě dat rozhodovat. Schopnosti a kulturu práce s daty vytváří celá řada lidí. Tato kniha vám bude praktickým rádcem, který vám na této cestě pomůže, ať už jste ředitel, nebo analytik. Plynulost nakládání s daty si žádá:

- **Vedoucího společnosti** – který chce založit kulturu předávání dat a práce s nimi a který si přeje vystavět tým důvtipných uživatelů a analytiků, již s daty pracují.
- **Nováčky** – kteří se učí, jak data správně předávat.
- **Zkušené analytiky** – kteří pracují s daty podrobně a potřebují svá zjištění efektivně předat dál.
- **Autory datových výstupů** – kteří navrhují a sestavují řídicí panely, reporty a na data bohaté aplikace, ale současně chtějí, aby jejich práce oslovila posluchače.
- **Vedoucí technology** – kteří mají na starosti správu dat a sestavování datových řešení a starají se o to, aby společnost měla z jejich práce maximální přínos.
- **Personalisty** – kteří musí vybírat a vyškolit kolegy pro práci s daty.

První kapitola podrobně rozebere rámec plynulého přenosu dat (Data Fluency Framework). Poté se s nabytými vědomostmi seznámíte v kapitole 2 s několika oblastmi, na které byste se měli zaměřit, pokud se rychle chcete dobrat dat, jež jsou pro vaši roli podstatná.

Kapitola	Vedoucí společnosti	Vedoucí technolog	Personalista	Analytik	Tvůrce obsahu	Nováček
3. Jak se společnosti potýkají s problémy Zjistíte, kterým problémům čelí vaše společnost	✓	✓				
4. Konzumentův průvodce daty Dovednosti potřebné pro efektivního konzumenta dat			✓			✓
5. Tvůrcův průvodce daty Dovednosti potřebné pro předávání dat		✓	✓	✓	✓	✓
6. Kultura plynulého přenosu dat Zakládání pravidel a standardů pro kulturu plynulého přenosu dat	✓		✓	✓	✓	
7. Ekosystém datových výstupů Tvorba prostředí pro plynulé a efektivní předvádění dat	✓	✓			✓	

Obrázek 2: Průvodce kapitolami

Poznámky

1. Joshua Porter, Principles of Product Design. Bokardo.com (bez data). Staženo z *<http://bokardo.com/principles-of-product-design/>*.

Problém poslední míle

V této kapitole:

- Informační doba: potřeba přenášet data plynule
- Plynulost dat: Využijte potenciál svých firemních dat
- Poznámky

*Voda, voda kolkolem a na troud dřevo schne,
voda, voda kolkolem a k pití kapka ne!*

*Samuel Taylor Coleridge
(přel. Josef Palivec)*

Když se na začátku 20. století v Texasu objevila ropa, populace Dallasu se během pár let několikrát znásobila. Tento obrovský růst hnala ekonomika Spojených států amerických, která přešla k masové výrobě, ale také infrastruktura, jež se přizpůsobila automobilům. První model Fordu T opustil továrnu 27. září roku 1908. Ani ne o dvacet let později, v roce 1927, měl Ford za sebou již 15 milionů vyrobených aut. Poptávka a nabídka automobilového průmyslu s sebou nesla obří společenskou změnu – Spojené státy americké doslova nastartovaly novou společnost.

Podobně i příchod a rozšířené používání osobních počítačů na konci 20. století stály na počátku dalšího boomu – věku informatiky. Třicet let po uvedení osobního počítače jsme vstoupili do éry veledat (big data) a rozhodování na základě dat, kterou přirovnáváme k ropnému boomu. Nadšení je rozhodně velké a doufáme, že data povedou k lepšímu porozumění a chytřejšímu podnikání.

Zní to slibně, ale má to háček. Celé generace vstupovaly rituálně do světa dospělých tak, že se naučily řídit, čímž se jim otevřel svobodnější svět. Učení se základním dovednostem pro práci s daty jsme si však doposud neosvojili. Jazyk, který stojí za hlavními obchodními tepnami, je nyní, a tím více i v budoucnu, dostupný pouze nemnohým.

Ve dvacátých a třicátých letech dvacátého století jsme byli svědky nástupu automobilek, které započaly s efektivní masovou výrobou. Již po generace získáváme z tohoto odvětví údaje, které nám pomáhají uvažovat o spotřebě ve vzorcích. Dnes, o téměř sto let později, potřebujeme získat základní dovednosti, které zajistí, aby jedinci i společnosti nakládali s daty s kritickým nadhledem. Tyto rodící se postupy jsou nezbytné, pokud chceme z dat sestavovat prezentace a použitelné výtvary.

Kniha je určena právě rozvoji těchto základních dovedností – informovanosti, schopným spotřebitelům dat, ale také tvůrcům smysluplných prezentací a výstupů. Naším cílem je pomoci jednotlivcům i firmám porozumět a rozvíjet plynulost předávání dat. Jak tvrdíme, data jsou novým jazykem, novou dálnicí obchodu 21. století.

Informační doba: potřeba přenášet data plynule

Neuvěřitelný pokrok ve schopnosti uchovávat data od základů změnil naše uvažování nad tím, co uchovat a co smazat. Nemusíme již pečlivě rozvažovat, které prvky digitální reality si ponecháme. Můžeme mít vše – a také tak činíme. Podle nezávislé výzkumné organizace SINTEF (The Foundation for Scientific and Industrial Research) vzniklo 90 procent všech dat na světě za poslední dva roky.¹ Díky četným obrázkům na Instagramu, tweetům a historii pohybu zákazníků na webech značně narůstá množství informací, které můžeme prosívat.

V jistém ohledu jsme nikdy dříve neměli více materiálu k tomu, abychom se mohli rozhodovat na základě informací. Ale přesto jsme stále daleci osvětlení. Studenti, učenci, zaměstnanci i vedoucí se stále často rozhodují náhodně, podle intuice a nálady davu. Pokud má být informační doba stejným boomem jako ropa, co představuje zisk? Rozhodujeme se lépe a rozumíme světu lépe? Jezdíme po digitální dálnici v autech, nebo ještě stále na koních?

V základu se dá říci, že ve firmách, s nimiž jsme ve společnosti Juice Analytics pracovali, byli rozladěni tím, jak nedokáží s daty pracovat efektivně. Tyto firmy si sice vystavěly datová úložiště, investovaly do drahých analytických nástrojů a utrácely své omezené zdroje na odborníky na data. Těžily data, analyzovaly je, definovaly metriky a sestavovaly přehledy. I přesto se v nich však data používala málo a se špatným porozuměním.

Jen málo lidí, a ještě méně společností, pracuje s daty konzistentně a nechává se jimi vést při rozhodování. Byli bychom rádi, kdyby celé té hory dat, které jsou k dispozici, využívali jak primární prodejci, tak i vrchní vedení firem. Zapomeňte na komplexní obchodní model Wall Streetu nebo na počítač společnosti IBM, který umí diagnostikovat nemoci. Data by měla být ve firmě používána jednoduše a postupně tak, aby posilovala diskuzi, zaměřovala zdroje a priority a také umožňovala provádět informovaná malá, každodenní rozhodnutí.

Řádné využití dat je problém, s nímž se potýkají malé i velké, státní i soukromé firmy napříč celým trhem. Podle studie, kterou vypracovala konzultantská firma Avanade, „více než 60 procent respondentů tvrdí, že se zaměstnanci musí naučit vytvářet z veledat (big data) hodnotu a obchodní přínos“.²

Data nesou obrovský příslib a mohou nám pomoci se mnohem lépe rozhodovat. Otázkou zůstává, proč je mezi nesmírnými možnostmi dat a skutečným stavem firem, které se s nimi potýkají, taková propast? Zde je několik teorií:

■ **Mnozí lidé se dat bojí a nechtějí se na jejich základě rozhodovat**

V jedné z předních internetových mediálních společností jsme byli svědky toho, jak analytický tým pečlivě vypracovává reporty, které ukazovaly, jak jsou jednotlivé online prvky úspěšné, zatímco jejich příjemci, vedoucí, předkládané údaje s pečlivostí o nic menší ignorovali. Rozhodování pak probíhalo intuitivně, a to i přesto, že měli vedoucí k dispozici více podrobných dat o využívání obsahu i o uživatelích.

■ **Technologická a personální omezení zabraňují společnosti pracovat se svými zdroji dat efektivně**

Mnohé společnosti, na něž jsme narazili, si stěžovaly na to, že jsou řízeny tabulkami. Každé oddělení má nějaká pravidla sběru, analýzy a reportování jedinečných dat. Neexistuje zde žádná jedna „pravda“ a datoví analytici se stávají nepostradatelní, protože jsou to jediní lidé ve firmě, kteří ví, jak fungují finanční modely, jak se dostat ke zdrojům dat a pracovat s nimi, ale znají také jejich silné a slabé stránky. Lidé si pak v takových společnostech přejí dobrat se technického řešení, které by dokázalo dát všechna data dohromady a zpřístupnit je všem, kteří se rozhodují, v jednoduchých, interaktivních přehledech.

■ **Organizační omezení znemožňují používat data efektivně**

Školští inspektoři po celé zemi si často povzdychují, že nemají k dispozici kvalitní data. Nadace Gates Foundation nyní vkládá značné prostředky do vývoje přehledů o jednotlivých krajích, na jejichž základě by bylo možné se rozhodovat. Nicméně s ohledem na včasný sběr dat, jejich propojení a na schopnost vyhodnocovat je a zapojovat do procesu výuky ještě stále zůstává řada překážek.

Jsme přesvědčeni, že strach z dat, technologická omezení, ale i nefunkční společnosti jsou pouhými symptomy něčeho rozsáhlejšího – nestojí za skutečností, že pro nás data nemají v současné době dostatečný přínos. Klíčovou příčinu spatřujeme v tom, co označujeme jako „problémem poslední míle“. V základu se jedná o to, že naše selhání při práci s daty není technologický, nýbrž společenský problém.

Připodobnění k poslední míli vychází z oblasti telekomunikací, kde je nejdražší přemstit posledních pár metrů mezi velkokapacitními sítěmi přenášejícími gigabajty dat a jednotlivými domy. Co se dat týče, sbírat a uchovávat je je snadné. Po technologické stránce máme již hotovo. Největší výzva teď čeká na analytiku, aplikace a jejich zapracování. Data lze sice sbírat hromadně, nicméně poslední míle je u nich vždy přizpůsobena dané společnosti. Pokud totiž chcete najít tu správnou perspektivu, řídit rozhodnutí a jednat, budete potřebovat dovednosti i nadšení na osobní i organizační úrovni. A to je ten chybějící článek – poslední míle, která si žádá, aby v práci s daty byl zblhlý jednatel i firma.

Tato kniha popisuje, jak mohou společnosti komunikovat s daty efektivněji – a to jak interně, tak i externě. Popisuje lidi a konkrétní dovednosti, které jsou nezbytné k tomu, abyste mohli data účelně vyhodnocovat a vytvářet na jejich základě reporty a prezentace.

Plynulost dat: Využijte potenciál svých firemních dat

Data se v mnoha ohledech podobají ropě – docela jistě v tom, že jsou ekonomickým motorem vaší společnosti. Surovou ropu nemůžete jen tak čerpat ze země přímo do nádrže auta nebo z ní vyrábět plastové kostičky LEGO. Podobně ani data nemůžete do společnosti jen tak numpumpovat a čekat, že vám budou nějak užitečná. Vztít si z dat přidanou hodnotu je složitý oříšek a jen nemnoho společností ho rozlouskne. Ačkoliv neexistuje jednoduché řešení (a proto se s daty tolik společností potýká), dobrou zprávou je, že když porozumíte povaze problému, budete mít již namířeno správným směrem. Plynulost přenosu dat je cestou – možností používat jazyk dat k plynulé výměně a průzkumu nápadů, které jsou pro společnost důležité.

V této knize se vám budeme snažit pomoci využít potenciál dat ve vaší firmě. Hlavními problémovými oblastmi pro vás nebudou technologie a omezení společnosti, ale rozvoj kapacit pro zpracování dat a jejich předávání mezi jednotlivci a týmy ve společnosti.

Plynulost předávání dat se týká jednotlivců (každý musí umět „číst a psát“ a „poslouchat a reprodukovat“ pomocí dat), ale také společností, které musí vytvářet prostředí, jež odmění produktivní diskuze o datech. Práci s daty se na poli jednotlivců věnuje řada praktických pramenů a knih. Nicméně již méně pramenů pomáhá přejít k plynulému převodu dat firmám. Tato kniha staví na základní literatuře popisující rozvoj společností, ale také na doporučeních nejúspěšnějších průmyslových gigantů současnosti.

Cílem je nabídnout vám základy, které by vám pomohly pochopit, co vše budete potřebovat k tomu, abyste ve své firmě dokázali přenášet data plynule. Rovněž vám nabídneme praktické rady, jimiž se můžete řídit. Nepohybujete se v teoretickém světě, takže se v něm tato kniha nebude zastavovat. Budeme se naopak zaměřovat na praktické příklady, které by vám měly pomoci načrtnout vlastní kroky.

Při radění budeme předpokládat, že nerozumíte technologické stránce věci. Existuje řada nástrojů, které vám mohou pomoci pracovat s daty plynuleji. Některé se vám budou líbit více a některé méně. Volbu mezi nimi často určí situace. Rady podané v této knize můžete použít na celé spektrum konkrétních produktů a platforem, s nimiž pracujete. Pomohou vám využít je co nejlépe.

Veledata a datová přeměna

Data jsou všude. Peta-, exa- a zettabajty dat. Jsou to množství, která si žádají neuvěřitelná přirovnání. Například 40 zettabajtů digitálních dat, které vyprodukuje do roku 2020, bude odhadem představovat 57násobek všech zrněk písku na všech plážích světa!³ Ačkoliv toto číslo až bere dech, data sbíráme stále z více pramenů a rychleji než dříve. Společně s nárůstem množství a typů dat přichází také technologická kapacita potřebná k jejich uchování a analýze.

Veledata (big data) – tzn. objem, rychlost a různorodost, s níž se dnes data shromažďují a uchovávají⁴ – představují budoucí hranici, a to nejen pro obchod, ale i pro další odvětví, včetně vlády, školství a zdravotnictví. Představitost i zájem ve vás jistě vyvolají příklady toho, jak vynalézavě lze data používat:

Statistická práce Billyho Beana, který se snažil týmu Oakland Athletics pomáhat drafťovat hráče pod cenou, zajistila týmu postup, ačkoliv na platy vydával tým pouhou třetinu toho co New York Yankees.

Analytické vzorce, které nasadil Google na Internetu, aby dokázal předvídat epidemie chřipky a horečky dengue.

Financování různých ministerstev federální vládou Spojených států amerických, které má „značně vylepšit nástroje a postupy, jež jsou zapotřebí při přístupu, organizaci a vytváření závěrů z velkoobjemových digitálních dat“.⁵

Profesionální technici, vedoucí hráči na poli obchodu, ale i akademici vnímají data jako nástroj sloužící k předvídání, snižování nákladů, vyhledávání obchodních příležitostí, vylepšování výzkumu a výkonnosti prodeje a výroby. Proto trh zaplavily startupy věnující se veledatům (big data). Investoři vkládají do tohoto sektoru ročně více než miliardu dolarů a sponzorují tak stovky společností.⁶

Skutečně se zdá, že jsou možnosti nekonečné. Představte si, že by všechna vaše data ve firmě zapadala do sebe, byla by kdykoliv k dispozici a pouze by čekala, až je zpeněžíte. Z tohoto hlediska působí data opravdu jako novodobá ropa. Pravdou je, že současnou éru veledat již popisuje řada přirovnání. Tyler Bell, analytik nakladatelství O'Reilly, jenž se věnuje veledatům, je třídí dle následující typologie:

- 1. Přirozené prameny** („novodobá ropa“, „zlatá horečka“ a samozřejmě „těžba dat“) – Zdůrazňují hodnotu, která je datům vlastní, s níž si pohráváme, když se snažíme využít jejich potenciál.
- 2. Přírodní katastrofy** („datové tornádo“, „záplava dat“, „datový příliv“) – Označuje data jako problém téměř biblických rozměrů, který se snadno přehoupne do katastrofy, když se včas a řádně nepřipravíme.
- 3. Průmyslová zařízení** („datový výfuk“, „hasičská hadice“, „průmyslová revoluce“) – Hodící se terminologie, která obvykle zobrazuje data jako mechanismus, který jsme si vytvořili a máme nad ním kontrolu, ale současně jako mechanismus, který se nám vymstí, pokud s ním nebudeme nakládat správně.⁷

Přirovnání, které bude nejvíce sedět na vaši firmu, bude samozřejmě dáno kapacitou společnosti. Patříte mezi 73 procent nedávno zkoumané skupiny amerických firem a vedoucích pracovníků v oblasti IT, kteří přiznávají, že převádět velké objemy dat do použitelných informací je pro ně stále výzvou?⁸ Anebo jste spíše jako polovina reprezentativního vzorku řídicích pracovníků na světě, kteří říkají, že mají k dispozici více dat, než s kolika dokáží pracovat? Chápete data jako surovou ropu, něco, co je třeba vytěžit, zpracovat, rafinovat, jak vzácný pramen, který je třeba upravit, pokud z něho máte získat nějaký přínos? Nebo spíše jako zlato, tedy něco, co je

třeba najít a vytěžit? Nebo možná jako hasicí hadici, která je nezbytná při hašení požáru, ale má příliš velký proud, když se chcete jen napít?

Uvedená přirovnání v sobě, jak vidíte, obsahují i postoj, který zaujímáte k datovému boomu – od bezbřehého nadšení až po zoufalá varování před použitím (a zneužitím) veledat, varování, která zdánlivě vychází jen z velkých úniků dat.

Někde mezi těmito dvěma extrémy se nachází většina lidí, jednotlivců, kteří vnímají data jako přínos, ale kteří jsou současně nešťastní, že jejich hodnotu nedokáží nijak zpeněžit. Dokonce i přední výzkumná společnost Gartner má pocit, že veledata začínají být zdrojem deziluze, protože zájem o ně opadá a společnosti neví kam dál, protože se jim nedaří získat z veledat přínos.⁹

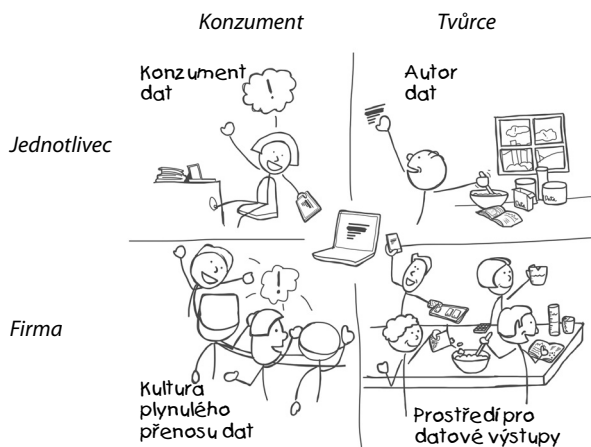
Co způsobilo úpadek? Proč přestávají mnozí doufat, že by data přinášela nové možnosti a přísliby?

Jednoduše řečeno si stále více lidí začíná uvědomovat, že samotný sběr dat, vývoj pokročilých úložných systémů a najímání odborníků nepředstavuje automaticky konkurenční výhodu. Důkazem pro toto tvrzení jsou snahy nesčetných vedoucích, zdravotnických firem, školních inspekci a dalších vládních i nevládních organizací. Studie společností, které začaly používat různé analytické nástroje, z roku 2013 ukazuje, že pouze 28 procent společností má pocit, že pro ně jejich investice měla citelný přínos.¹⁰ Rozladění z dat pak pramení nejen z nenaplněných očekávání, ale také ze skutečnosti, že jejich koncoví uživatelé nedokázali data smysluplně zpracovávat a předávat dále. Tato kniha se pokusí podat na tento neduh lék, a to jak na úrovni jednotlivců, tak i na úrovni celé firmy.

Náš rámec pro plynulý přenos dat

Ve snaze pomoci vám využít potenciál dat při práci jednotlivců, v týmech i celých společnostech jsme vyvinuli rámec pro plynulý přenos dat. Rámec, který vidíte na obrázku 1.1, zobrazuje potřebné dovednosti a kompetence, které můžete rozvíjet.

Rámec definuje dvě základní kategorie dovedností, které jsou pro rozvoj plynulého přenosu dat nezbytné. Konkrétně se jedná o schopnosti, které potřebuje znalý konzument prezentací obsahujících data, ale také o odbornost, již potřebují vyškolení tvůrci těchto prezentací. Rozvoj i aplikace těchto dovedností se projevují na dvou úrovních, na individuální i firemní. Cílem je urychlit plynulé přenášení dat tak, jak to vidíte ve všech čtyřech kvadrantech rámce.



Obrázek 1.1: Rámec plynulého přenosu dat

Případové studie: Okno do rámce plynulosti předávání dat

Další kapitoly této knihy mají za úkol pomoci jednotlivcům i společnostem rozvíjet dovednosti znázorněné ve všech kvadrantech rámce plynulosti přenosu dat. Tento oddíl nabídne čtyři případové studie a ukáže vám tak náhled rámce a odhalí typ popisovaných dovedností. Ke každému problému či doporučení, které je v případech načrtnuto, nabídnou následující kapitoly další rady, které umožní růst jak jednotlivcům, tak firmám.

Konzumenti dat: Fantasy Football

Nejspíš to nebude ani nadsázka, když řeknu, že se v Americe naučí datové statistice více lidí pomocí hry Fantasy football (statistická hra na manažera týmu amerického fotbalu, pozn. překl.) než ve statistických seminářích. Každou neděli se snaží 19 milionů fanoušků národní fotbalové ligy odhadnout výkon svého týmu na základě statických předpokladů, historických vzorců a analýzy týdenních odchylek. Povaleči, kteří se dříve bavili na hřišti, nyní tráví průměrně více než osm hodin týdně ponoření do sportovních dat.

Pokud tuto hru neznáte, Fantasy football umožňuje fanouškům zahrát si na vlastníka a manažera vlastního smyšleného týmu a každý týden hráče nově rozestavit. Zatímco se pak každý týden odehrávají zápasy, pomyslní vlastníci sbírají body a ve smyšlených ligách soupeří s ostatními hráči. Pokud chtějí vyhrát, brzy si musí všimnout, že úspěch často závisí na podrobném studiu výkonů hráčů a týmu.

Níže najdete několik způsobů, kterými hráči hry Fantasy football zapracovávají data do vlastního uvažování:

Proměnlivý výkon hráče

Nejlepší vlastníci ví, že se výkony hráčů každý týden mění, a chápou, jak jsou svázány s bodováním. Kupříkladu touchdown přinese vlastníkovvi běžně šest bodů. Jenže touchdown se podaří málokdy a četnost jeho výskytů může divoce kolísat. Naopak počet získaných míčů může být lepším ukazatelem toho, jak tým pracuje a kolik bodů může majiteli přinést. Je však důležité, aby byl výkon vyvážený, takže úspěšní vlastníci se více soustředí na hráče, u nichž lze předvídat vyváženější výkon (kupříkladu na základě počtu získaných míčů), nikoliv na hráče s vyšším počtem náhodných výkonů (například touchdownů).

Hodnocení může mást

Pomocné nástroje nabízí hodnocení hráčů na všech pozicích. Hodnocení však skrývají rozdíly a rozptýlení očekávaného výkonu. Kupříkladu od nejrychlejšího backa se může očekávat, že bude o 20 % lepší než druhý nejlépe hodnocený back, od kterého se dále může čekat, že dá jen o pět procent více bodů než třetí až šestý nejrychlejší back. Data ukazují, že se hráči často shlukují do výkonnostních bloků. Popsané chápání statistiky veřejně popsal Boris Chen, který prohlásil: „Hráči v daném bloku jsou si víceméně rovni. Míra šumu mezi jednotlivými pořadími v daném bloku a skutečnými výsledky je natolik významná, že většinou si můžete v podstatě hodit kostkou.“ Tento princip pak vlastníci smyšlených týmů povětšinou přijali a hráče do týmů si vybírali v souladu s ním.

Jedinou konstantou je změna

Nejhorší hráči se upínají k minulosti a volí si hráče a týmy podle toho, které týmy jim přinášely body již dříve. To znamená, že se nedokáží přizpůsobit novým výsledkům týmů a hráčů ani trendům. Když sledujete data pečlivě, poznáte, kdy hráč dosáhl vrcholu a kdy tým, který dříve s přehledem vítězil, již nedokáže zaskórovat. Snaha držet se zažitých pořádků může být cestou ke katastrofě, protože jedinou konstantou hry je změna.

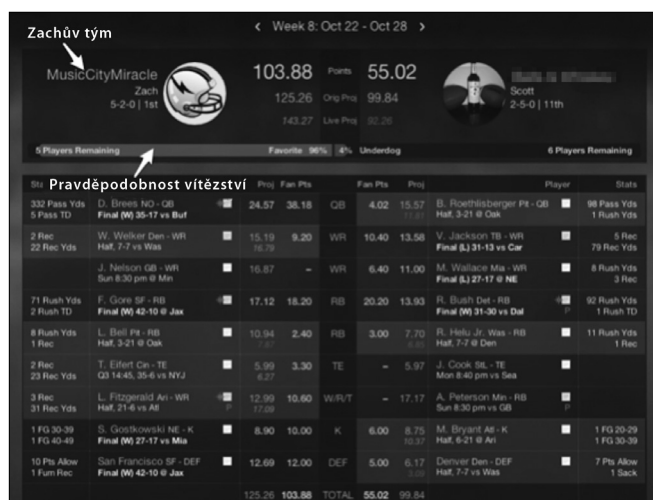
Kontext vám dokreslí obrázek

Data nahlížená bez kontextu mohou mást. Řekněme například, že váš nejlepší wide receiver dal za sezónu pouze polovinu bodů, které běžně zvládne. Jedná se o nový, znepokojující trend? Měli byste ho prodat? Když se trochu zahloubáte do dat, zjistíte, že hrál proti jednomu z nejlepších cornerbacků ligy anebo že mu ze hry vypadl quarterback či že špatných výsledků dosahuje při špatném počasí a na cizích hřištích. S ohledem na výsledky mohou být tyto vlivy prostředí významné. Výkonnostní data nelze zkoumat izolovaně – kontext je důležitý.

Jak se tedy smyšlené lize amerického fotbalu podařilo sestavit legie fanoušků, kteří si vytvořili vlastní dialekt, aby si mohli data předávat plynule? Může za to propojení učení, efektivní prezentace dat, běžných zvyků při práci s daty a pobídek. Manažeři fiktivních fotbalových klubů se učí pracovat s daty ve svůj prospěch pomocí NFL, ESPN, Yahoo! a řady dalších webových stránek, které se věnují analýze amerického fotbalu. Společnosti, mezi něž patří Football Outsiders, vystavěly na modelování dat a předvídání výkonů hráčů nová mediální odvětví.

Přední stránky zprostředkovávající fiktivní fotbalovou ligu, jako například ESPN a Yahoo!, data a jejich vizualizace uživatelům agresivně podsouvají. Na těchto stránkách najdete grafy popisující vývoj jednotlivých hráčů, grafy nasazení hráčů, grafiku srovnávající hráče a prediktivní modely, které předvídají výsledky hry.

Výškolená komunita kolem fiktivního amerického fotbalu se o sport vášnivě zajímá. Hráči americký fotbal milují! Fiktivní liga nabízí svým fanouškům zcela nový (a vděčný) rozměr. Neomezuje ji již fandění jednomu týmu – do zorného pole hráčova zájmu se dostává celá liga, v níž si může volit ze všech hráčů 32 týmů NFL. Odvětví fiktivního amerického fotbalu se navíc drží běžných formátů ligy, bodování i klíčových metrik (obrázek 1.2). Pojmy, mezi něž patří PPR, RBBC, waiver wire (systém najímání hráčů, pozn. překl.) a flex (flexibilní pozice, pozn. překl.), jsou dobře srozumitelné a hráčům fiktivní ligy usnadňují komunikaci. S ohledem na 1,18 miliardy dolarů, které hráči ročně ve fiktivní lize prosází, a vášnivou základnu fanoušků mají hráči velkou motivaci rozhodovat se správně. Když jde o peníze a čest, lidé investují do studia plynulého přenášení dat čas i energii.



Obrázek 1.2: Stránka Yahoo! zobrazující stav fiktivní ligy amerického fotbalu

Jednoduše řečeno, uvedené faktory předložily plynulý přenos dat masám. Miliony fanoušků se naučily číst grafy, porozumět základním datovým principům, ale také se naučily rozhodovat se na základě dat, která nejsou viditelná na první pohled – to vše jsou základní dovednosti spadající do prvního kvadrantu našeho rámce.

Tvůrci datových produktů: U.S. News

Rektorka univerzity, která se v celonárodním hodnocení umísťuje na předních příčkách, čeká na zprávy. V minulém roce se škola na seznamu nejlepších vysokých škol umístila dle hodnocení mediální společnosti U.S. News & World Report na dvacáté příčce a s ohledem na početnou a vlivnou základnu absolventů byla rektorka pod velkým tlakem. Byla sice připravena

prohlásit, že na hodnocení nezáleží, ale rozpomínala se na to, jak se na každé větší náborové a fundraisingové události umístěním na dvacáté příčce chlubila. Nedaleko od její kanceláře se budou studenti končících ročníků na nejlepší střední škole ve městě rozhodovat, na kterou vysokou školu půjdou.

Jde o hodně. Ekonomové James Monks a Ronald Ehrenberg zjistili, že když si škola v hodnocení pohorší, sníží se počet podaných žádostí o přijetí, kvalita uchazečů, ale také množství sponzorských darů (Monks a Ehrenberg, 1999)¹¹. Nedávno provedená studie dále pak zjišťuje, že posun nahoru s sebou přináší značný nárůst přihlášek, a to napříč všemi typy institucí (Bowman a Bastedo, 2009)¹².

Hodnocení je ale přesto zdrojem kontroverze. Jeho kritici tvrdí, že hodnocení příliš neodráží změny provedené v daném roce – čímž se hodí spíše těm, kteří jsou setrvale nahoře, například Harvardu, Princetonu, Stanfordu a hrstce dalších prestižních škol. Zastánci však oponují, že hodnocení pomáhá vyhodnocovat zavedené školy a ukazuje veřejnosti, jak si školy stojí v oblasti hospodaření, kvality vyučujících, spokojenosti studentů a množství dalších ukazatelů. Rozhodnout se mezi 2 400 státními a soukromými školami nabízejícími čtyřleté obory není snadné.

I přes rozporuplné pocity, které hodnocení jitrí, oblíbenost hodnocení roste. Kupříkladu vždy v září, kdy společnost U.S. News vydává seznam nejlepších vysokých škol Best Colleges, přichází na webovou stránku USNews.com obrovské množství návštěvníků. V roce 2013 se na stránku přišlo v době vydání hodnocení podívat 2,6 milionů jedinečných návštěvníků, kteří stránku zobrazili 19milionkrát.¹³ Zájem na Internetu ukazuje, že lidé považují hodnocení za neuvěřitelně praktický nástroj, který jim umožňuje prosívat obrovská množství informací, jež jsou o školách k dispozici – informace, které každá škola prezentuje jinak – a které jim pomáhají si vybrat vhodnou školu.

Co se tedy od společnosti U.S. News, která předává data širokému publiku, můžeme naučit? Jak je možné, že se této společnosti podařilo vytvořit tak úspěšný formát? Jak se z ní stal neoficiální průvodce po vysokých školách?

Seznam hodnocených škol Best Colleges společnosti U.S. News začínal původně jako průzkum oblíbenosti rektorů. Časopis People měl svůj seznam 50 nejhezčích lidí, Forbes seznam 400 nejbohatších lidí. Oba časopisy patřily k těm nejprodávanějším. U.S. News a jeho vedení dobře chápe moc hodnocení. Bob Morse, hlavní stratég pro data ve společnosti U.S. News, si spolu s kolegy uvědomoval víc než jen to, že seznamy prodávají, dokázali data také dobře sestavovat. Uvědomili si, že hodnotící seznam je mocný nástroj, který dokáže seskupit obrovská kvanta informací do jednoduché, snadno stravitelné tabulky.

Při sestavování a šíření hodnocení nejlepších vysokých škol si Morse se svým týmem počínají v souladu se čtyřmi důležitými zásadami, jichž se drží každý zkušený tvůrce dat.

Kategorizace informací

Morse roztrídil instituce nabízející vyšší vzdělání do kategorie na základě Carnegieho klasifikace (Carnegie Classification of Institutions of Higher Education). V průběhu času si pak společnost U.S. News přidala kategorie vlastní, které splňovaly nároky konzumentů. Americké

vysoké školy a univerzity například třídí na státní univerzity, státní vysoké školy svobodných umění, oblastní univerzity, oblastní vysoké školy, praktické školy, pokrokové školy (up-and-coming schools), nejlepší státní školy a školy s nejvyšším počtem zahraničních studentů. Tyto kategorie podávají studentům do začátku hledání důležité informace.

Důraz na hodnocení a srovnání

Když už si tvůrci hodnocení založili kategorie, mohou v nich na základě dalších informací jednotlivé informace porovnávat a řadit. Ke srovnání mohou posloužit například proměnné, jakými jsou školné, procentuální úspěšnost uchazečů, průměrná délka setrvání přijatých studentů, poměr absolventů, velikost tříd a průměrné výsledky v testech SAT. Prostřednictvím srovnání a hodnocení se data prezentují velmi jednoduše.

Snaha pomoci konzumentům data pochopit

Do hodnocení se postupně dostaly různé proměnné. Kupříkladu v roce 1983 přibýlo hodnocení inaugurací, které se zakládalo na průzkumu, v němž mělo 1 308 rektorů vysokých škol a univerzit zvolit pět nejlepších škol ve své kategorii. Od roku 1988 byly součástí hodnocení i údaje o výběru uchazečů, postavení fakulty, prostředcích vydávaných na vzdělávání a počtech absolventů.¹⁴ V průběhu času tak Morse se svým týmem zapracovával do modelů pro výpočty hodnocení nové proměnné. Jakmile si veřejnost osvojila základy, mohla společnost U.S. News hodnocení dále rozšiřovat.

Zpřístupnění dat různým příjemcům

Ve snaze pomoci studentům a jejich rodinám při výběru vysoké školy rozjela společnost U.S. News nástroj pro výběr školy (My Fit Engine), který na základě zadaných preferencí zobrazí odpovídající školy. Ačkoliv může hodnocení pomoci studentům a jejich rodinám probrat se vysokými školami, stejně tak (ne-li více) pomáhá i vysokým školám. Při průzkumu mínění na vysokých školách společnost U.S. News zjistila, že se školy často zamýšlejí nad minulými trendy a hledají v nich vzory. Podobným způsobem pracovala společnost U.S. News se společností Juice Analytics, když vytvářela online nástroj Academic Insights, který analyzuje a zkoumá data v průběhu času.

U.S. News přináší značnou inovaci v tom, že obrovská množství informací vkládá do jednotného formátu, který je snadno srozumitelný, a ten pak takto zpracované údaje prezentuje v praktické tištěné a online podobě. Společnost U.S. News vytvořila systém, ve kterém lze v několika seznamech zobrazit údaje o stovkách institucí. Rodina se tak nyní již nemusí probírat nespočetnými webovými stránkami a reklamními brožurami. Množství času stráveného výběrem vysoké školy si dnes může ušetřit. V tomto ohledu představuje U.S. News dovednost, která v našem rámci dominuje druhému kvadrantu – schopnosti prezentovat data účelně.

Konzumenti ve firmách: inspekce hořekuje

Schůzka nedopadla dobře. Školská inspekce se už pět let snaží zapracovat nový systém hodnocení učitelů – a s ním i odpovídající databázi a reportovací systém, který by s ní umožňoval

pracovat. Na nábor „datových školitelů“, kteří měli ředitele a učitele seznámit s datovými systémy a naučit je s nimi provádět základní operace, inspekce dokonce vyčlenila značně vysoké procento ze svého desetimilionového grantu, jež získala od federální vlády. Jenže peníze poslední rok došly a školitelé museli skončit, mnozí odešli jinam. Ta hrstka těch, kteří zůstali, se sice uchýlila na inspektorát, kde měla s IT oddělením spolupracovat na široké škále projektů, ale neměla již žádný společný cíl.

Nyní, o rok později, svolala Školská rada zvláštní schůzi, jejímž cílem bylo zjistit stav a dopad tohoto drahého projektu. Z centrály přišla hrozivá zpráva – školitelé již nemohli pracovat ve školách a bez dostatečných financí, které byly třeba k dalšímu profesionálnímu rozvoji, přestali ředitelé i učitelé nejspíše systém používat. Členové rady, kteří chtěli vědět, jak systém pracuje, vyvíjeli na inspekci nátlak a tázali se například: Jde o technický problém? Nemá váš tým dostatek potřebných pracovníků, kteří by se o chod projektu starali? Změnilo se s výjimkou školitelů od minulého roku, kdy se zdálo, že systém používají všichni, ještě i něco jiného? Inspekční tým odpověděl, že datový systém nejenže fungoval bezchybně, ale že ho i vylepšili tak, aby měli ředitelé a učitelé přístup k více reportům a nástrojům, které by jim pomohly prohlížet a vizualizovat data.

„Dnes již spolupracujeme s IT oddělením, a tak je systém mnohem lepší,“ prohlásil ředitel oddělení. „To, co v systému mohou učitelé dělat, neustále vylepšujeme – reporty i nástroje. Každý týden posíláme e-mail, v němž informujeme o všech změnách a ukazujeme reporty, které jsou dle našeho mínění nejužitečnější. Zpětnou vazbu máme ale pouze od několika lidí, od nadšenců do technologií, kteří mají data rádi. Nevíme, proč ostatní systém nepoužívají. Je zřejmé, že jim na systému něco vadí. Ale neřeknou nám, co!“

V rámci prezentace představil tým i údaje z průzkumu, který prováděl v daném kraji na náhodném vzorku škol. Výsledky nebyly povzbudivé. Méně než 10 procent respondentů uvedlo, že systém používá „pravidelně“, a přibližně 55 procent vybralo možnost „nikdy“. Následující otázky, které zjišťovaly, proč uživatelé systém nepoužívají, ukazoval, že značná většina učitelů a ředitelů má pocit, že je datový systém s reporty „příliš složitý“. Navíc měli pocit, že je kraj „zahltl informacemi“, ale pouze 25 procent mělo pocit, že šlo o „podstatné informace“.

Po prezentaci výsledků výzkumu dostala skupinka učitelů a ředitelů, kteří byli pozváni na schůzku, čas na diskuzi a odpověď na otázky. Jejich odpovědi vnesly do případu světlo:

Jeden z ředitelů řekl: „Nemám čas otvírat každý e-mail, který z datového oddělení dostanu. Když už ho otevřu, tak má pokaždé jiný formát. Někdy je to PDF, někdy ho musím otevřít v Excelu.“

Jeden z učitelů odpověděl: „Polovina z reportů, které dostáváme, není pro naši školu a studenty vůbec relevantní.“

Jiný učitel se tázal: „Jak mám pochopit, jak mi mají údaje pomoci, když ani nevím, jak je vypočítávají?“

Jeden z ředitelů dodal: „Pořád něco mění. Někdy se tam podívám a najdu to, co potřebuji, ale pak se na stejné místo podívám za několik dní a už to tam není! A navíc infor-

mace, které od nich dostávám, se liší od těch, které mám od personálního a od našeho školského inspektora.“

Jiný ředitel poznamenal: „Nemáme čas se zastavit a jako vedení se na data podívat.“

Poslední ředitel řekl: „Coby „vyškolení“ máme problém všem těm datům vůbec porozumět... Rodiče tomu nebudou rozumět už vůbec.“

Poté co všichni učitelé a ředitelé otázky okomentovali, strávila Školská rada zbytek schůze dohadováním se nad zásadní otázkou: Proč jsme investovali tolik peněz do systému, který se nepoužívá... A co budeme dělat dále?

Všechny tyto problémy budeme podrobně řešit v kapitole 6. Najdeme je ve všech organizacích, které se snaží položit základy systému, jenž by prezentoval data všem.

Tvůrci ve firmách: Limity pojišťoven

Velké společnosti utrácejí miliony dolarů ročně na krytí pojistných událostí zaměstnanců. Zdravotní pojišťovny pak musí vysvětlovat, kde všechny tyto peníze končí a proč se platby každoročně zvyšují, ačkoliv běží pečlivě vypracované programy, které mají zlepšovat zdraví pracujících. Všichni účetní se v pojišťovnách potýkají se stejným problémem: Jak vše sdělit zákazníkům pomocí dat?

Cílem je ukázat, že kvalita poskytovaných zdravotnických služeb roste, což přináší lepší zdravotní podmínky zaměstnancům a nižší náklady pro zaměstnavatele. Pokud účetní předá tyto informace správně, zaměstnavatel s pojišťovnou prodlouží smlouvu. Prezentace těchto údajů jsou však často spíše nudné, posluchače nezaujímou. Nedokáží podat pořádný příběh, a to proto, že ukazují na sobě nezávislá čísla, která pochází z několika databází. Na sběr dat, jejich srovnání na celostátní, oblastní a oborové úrovni bývá vynakládán značný čas, ovšem málo energie se vydá na příběh, který by data propojoval. Anebo opačně: Příliš mnoho energie padne na dobrý příběh, ale ten pak nebývá zaštitěn dostatečným množstvím dat. V obou případech po chabých prezentacích přichází požadavky na jiné metriky a více dat, což může, ale taky nemusí být k užtku. A pro ředitele účetního úseku, který se snaží uzavřít smlouvu, to může znamenat, že zákazník tentokrát smlouvu neprodlouží.

Představte si situaci, kdy musíte o čtvrt milionu zaměstnanců obří firmy zjistit údaje o nemocnosti a ty pak propojit s jejich výkonem, zdravím a délkou setrvání v práci, a to na místní i oblastní úrovni. Zaprvé, data o čtvrt milionu zaměstnanců budete sbírat z mnoha pramenů (zaměstnanec, zaměstnavatel, ale také různé zdravotnické a pečovatelské prameny: nemocnice, kliniky, poradenská centra, terapeuti, protikuřácké programy, online programy, a dokonce posilovny). Některé části programu se mohou konat přímo ve firmě, zatímco další, jako například odvykání kouření a boj s depresí, se budou odehrávat jinde. Posbíraná data o čtvrt milionu zaměstnanců poté spojíte s dalšími 40 miliony dalších životů, které pojišťuje tato velká pojišťovna – což je již velký datový fond, z něž lze již statisticky srovnávat ty, kteří se účastní zdravotních programů, a ty, kteří nikoliv. Každý den přichází nové údaje – někdy v noci v rámci synchronizace se zaměstnavatelovou databází zaměstnanců a zabezpečenými servery pojišťovací společnosti, ale někdy také poštou, to v případě žádostí jednotlivých zaměstnanců

o úhradu nákladů na posilovnu. Různá oddělení sbírají různá data do 52 různých databází. Vytáhnout, analyzovat a představit tato data zpět zaměstnavateli v nějaké srozumitelné podobě není snadné – musíte ukázat nižší náklady na zdravotní péči, zvýšit výkonnost zaměstnavatele a prodloužit dobu, po kterou v zaměstnání setrvá, ale také posílit zdraví a pracovní morálku zaměstnanců. Nárůst v těchto třech oblastech je nutno dát do kontextu na místní, regionální i celostátní úrovni.

Ačkoliv se zdá, že se jedná o příliš složitý úkol, velké zdravotní pojišťovny mají často jedny z nejsložitějších a neúplnějších databází zachycujících pojistné události zaměstnanců a historické trendy. Ve spojení se špičkovými týmy statistiků a datových odborníků by se zdálo, že poskládat příslušná data do prezentace pro velkého klienta bude jednoduché.

V datovém prostředí se však nachází spousta slabin, které brání systematickému rozvoji a sdílení prezentací, takže ve výsledku nelze zachytit celý příběh. Možná vám budou čtyři následující problémy připadat povědomé:

- Data se ukládají do mnoha oddělených, nesouvisejících databází. Data se často sbírají z desítek nezávislých sestav. Abyste sestavu dostali, musíte si poslat „ticket“, který vás u správce databáze zařadí do fronty čekatelů. Když data obdržíte, začnete uvažovat trochu jinak, takže si pošlete další „ticket“. A i kdyby byly sestavy k dispozici i bez správce databáze, strávili byste značný čas přihlašování se do různých datových systémů a sběrem potřebných údajů.
- Při propojování všech takových dat do jedné powerpointové prezentace je třeba postupovat opatrně. Dát data dohromady tak, aby působila čistě, profesionálně a lahodila oku, není tak úplně jednoduché. Různé sestavy vychází z různých metrik. Prezentovat data dostatečně jednoznačně, a navíc líbivým způsobem si žádá zkušenosti.
- Vyprávět příběh je tvůrčí akt. Zapojit metriky do příběhu, ne pouze do řady faktů, si žádá čas i úsilí. Navíc k tomu potřebujete lidi, kteří umí pracovat jak technicky, tak i kreativně.
- Najít průsečík mezi klientovým požadavkem na podrobný rozbor různých oblastí a nedostatečným časem není snadné. Když budete chtít vyhovět klientovu protichůdnému požadavku prezentovat podrobné údaje v 90minutovém sezení, budete muset přednést pouze nejdůležitější údaje – ale velmi přesvědčivě, povrchní přístup vám zde nepomůže.

Vyprávění přesvědčivého příběhu prostřednictvím účelně použitých údajů připravuje půdu novému druhu schůzek mezi společnostmi a klienty – schůzek, na kterých klient dokonce vůbec neuvažuje, zda se společností prodlouží smlouvu, ale spíše se snaží hledat řešení na problémy, o nichž se dozví z duchaplné analýzy. Konečně mnohé společnosti mají problém vyprávět příběhy založené na pečlivě připravených datových sestavách konzumentům, kteří nejsou po datové stránce zběhlí. Tyto dovednosti ze čtvrtého kvadrantu, které budeme rozebírat v kapitole 7, jsou nezbytné pro rozvoj prostředí, ve kterém lze data sestavovat i sdílet.

Poznámky

1. Dragland, Åse, „Big Data – for better or worse.“ SINTEF, použito 22. července 2013, <http://www.sintef.no/home/corporate-news/Big-data—for-better-or-for-worse/>.
2. Avanade, „Global Survey: Is Big Data Producing Big Returns?“ Publikováno v červenci 2012, použito 22. července 2014, <http://www.avanade.com/~media/documents/research%20and%20insights/avanade-big-data-executive-summary-2012.ashx>.
3. Lucas Mearian, „By 2020, there will be 5,200 GB of data for every person on the Earth,“ Computer World, použito 11. prosince 2012, http://www.computerworld.com/s/article/9234563/By_2020_there_will_be_5_200_GB_of_data_for_every_person_on_Earth.
4. Doug Laney, „3E data management: Controlling data volume, velocity and variety,“ META Group Research Note 6 (2001), použito 22. prosince 2014, <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>.
5. Office of Science and Technology Policy, „Obama Administration unveils ‚Big Data‘ initiative: Announces \$200 Million in R&D Investments,“ použito 22. července 2014, http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf.
6. Eric Sherman, „VC Bets \$1.4 Billion on Big Data in 2012,“ Inc., 28. únor 2013, použito 22. července 2014, [http://www.inc.com/erik-%20sherman/vcs-bet-\\$14-billion-on-big-data-in-2012.html](http://www.inc.com/erik-%20sherman/vcs-bet-$14-billion-on-big-data-in-2012.html).
7. Tyler Bell, „Big Data: An Opportunity in Search of a Metaphor,“ Radar, 10. února 2011, <http://radar.oreilly.com/2011/02/big-data-metaphor.html>.
8. CopTIA, „2nd Annual Big Data Insights and Opportunities,“ studie, použito 22. července 2014, <https://www.comptia.org/resources/2nd-annual-big-data-insights-and-opportunities>.
9. Svetlana Sicularová, „Big Data Is Falling into the Trough of Disillusionment,“ publikováno 33. ledna 2013, použito 22. července, <http://blogs.gartner.com/svetlana-sicular/big-data-is-falling-into-the-trough-of-disillusionment/>.
10. Cindi Howsonová, „BI Adoption Remains Flat“ BI Scorecard, 1. duben 2014, použito 22. července, <http://www.biscorecard.com/bi-adoption-flat/>.
11. J. Monks a R. G. Ehrenberg (1999). The Impact Of U.S. News & World Report College Rankings On Admissions Outcomes And Pricing Policies At Selective Private Institutions (CHERI Working Paper #1). Použito 1. června 2014, webová stránka ILR School, Cornell University, <http://digitalcommons.ilr.cornell.edu/cheri/1>.
12. Nicholas Bowman a Michael Bastedo, „Getting on the Front Page: Organizational Reputation, Status Signals, and the Impact of U.S. News and World Report on Student Decisions,“ Research in Higher Education, srpen 2009, 50. svazek, 5. vydání, str. 415–436. Použito 1. června 2014, <http://www-personal.umich.edu/~bastedo/papers/BowmanBastedo.ResHE2009.pdf>