

# KDO ANALYZUJE DATA, VIDÍ DÁL

Obchodní rozhodnutí, která jsou založena na datové analytice, mohou přinést rozhodující konkurenční výhodu, říká Jiří Bohuslav a Pavel Pilař ze společnosti Sophia Solutions.

**Základním kamenem úspěchu budou patrně zdrojová data. Vědí vůbec vaši klienti, jaká data mají sbírat?**

**Jiří Bohuslav (JB):** Zákazníci to většinou vědí a data mají k dispozici.

Problém je, že jsou uložena v různých systémech a mají různou kvalitu. Proto je zpravidla prvním krokem vytvoření datového skladu, do něhož periodicky přicházejí data z primárních systémů a dochází k jejich kontrole.

**Pavel Pilař (PP):** Často je zde i jiný problém: data se sice sbírají, ale firmy nevyužívají jejich potenciál. Zpravidla nejsou schopny efektivně propojit vlastní data s veřejnými zdroji, jako jsou údaje Českého statistického úřadu, etalon adres, informace o místním nákupním chování (geosoring) nebo kurzovní listek. Řada systémů má své reportingové prostředky, ale my jsme schopni data z různých systémů propojit, a dát tak firemnímu managementu mnohem širší pohled na celek. Výsledkem může být výrazná konkurenční výhoda.

**Můžete uvést nějaký příklad?**

**JB:** Velmi zajímavou úlohou je třeba odhalování pojistných podvodů. Například řetězový podvod u autopojištění. Ten vypadá tak, že se na něm podílí více osob, které jsou střídavě v roli viníka nebo poškozeného, a spojovacím článkem může být třeba autoservis, kde havarovaná vozidla končí. Pro likvidátora je velmi těžké takový podvod odhalit, protože vidí každou škodní událost samostatně – a ta sama o sobě nevzbuzuje pochybnosti. Teprve širší pohled na data může

vést k odhalení možného propojení a celý řetěz pojistných událostí může být označen jako podezřelý.

**Dovedli byste třeba podle mého dosavadního nákupního chování odhadnout, co si pořídím jako další věc?**

**PP:** I takovéto úlohy se běžně řeší. Zde už se dostáváme do ještě zajímavější ob-



**Jiří Bohuslav,  
Managing Consultant  
a Team Leader DWH,  
Partner Sophia Solutions**

Vystudoval Západočeskou univerzitu v Plzni, Fakultu aplikovaných věd, obor finanční a informační management. V oblasti datových skladů a reportingu se pohybuje již dlouhých 17 let, působil na mnoha projektech v rolích konzultanta, analytika, architekta řešení i projektového manažera. Do společnosti Sophia Solutions nastoupil v roce 2004, od roku 2012 je jedním z partnerů společnosti. Dlouhodobě vede tým specializující se na problematiku datových skladů a big dat.

lasti a tou jsou prediktivní modely. Firma chtěla prodat daný počet kusů určitého výrobku a hledala cestu, jak co neefektivněji tento výrobek svým zákazníkům nabídnout. Vytvořili jsme prediktivní model, který každému zákazníkovi přidělil skóre odpovídající pravděpodobnosti, že na danou nabídku zareaguje kladně. Navíc jsme byli schopni odhadnout, že stačí oslovit 30 % zákazníků s nejvyšším skóre. Pokud by firma oslovovala potenciální kupce náhodně, musela by nabídku směřovat na 50 % z nich. Rozdíl mezi těmito dvěma údaji je základem úspory, kterou naše řešení přineslo. Pokud prodejce využívá nějaký dražší obchodní kanál, jako jsou direct mailing, telemarketing nebo kombinace obojího, může být v takových případech úspora značná a navíc přesně vyčíslitelná.

**Takový model je asi nutné udělat vždy znova pro každý typ výrobku...**

**JB:** Ano, ale vytvoření každého dalšího modelu může být mnohem snazší. Data o zákaznících a o jejich dosavadním nákupním chování jsou již k dispozici a stačí je případně aktualizovat.

**PP:** Úlohu je možné také otočit – lze navrhnout model, který určí, jaké typy výrobků je efektivní danému zákazníkovi nabízet. Pod typem je možné rozumět i obecnou charakteristiku, třeba jestli kupuje značkové výrobky nebo zda preferuje nižší cenu.

**To připomíná klasickou segmentaci zákazníků, ovšem postavenou na mnohem komplikovanějších datech, než jsou jen pohlaví nebo věk.**

**PP:** Segmentační modely se v analýze často využívají a neslouží jen k marketingovým účelům, jak by se z předchozího příkladu mohlo zdát. Například jsme pomocí segmentačního modelu řešili odhalování podvodného chování vůči zdravotní pojišťovně. Zde se takového jednání mohou dopouštět poskytovatelé péče tím, že vykazují jiné, dražší výkony, než ve skutečnosti dělají. Model například odhalil, že struktura vykazované péče neodpovídá segmentu pacientů,

kteřé dané zdravotnické zařízení obsluhuje. Toto zjištění pak slouží jako podnět pro zahájení detailnější kontroly.

### Uplatňují se prediktivní modely i v jiných oblastech?

**JB:** Rozhodně, používají se třeba v oblasti prediktivní údržby ve výrobě. Příkladem může být výrobní linka na vstřikování plastů. Zákazník má moderní linku, která byla vybavena vstřikovacími lisami se spoustou senzorů. Prvotním cílem byl přehled o produkci a ukládat data pro podrobnou dokumentaci výroby. Během provozu se podařilo vysledovat, že některé kombinace údajů ze senzorů, jako teplota nebo tlak, mohou vést k poruše stroje a zastavení linky. Na základě toho je možné vytvořit prediktivní model, který sleduje senzory, vyhodnocuje údaje, a hlídá tak riziko poruchy. Obsluha může včas reagovat a upravit technologické parametry tak, aby se předešlo prostojům nebo úplné havárii stroje.

### Můžete jmenovat nějakou zajímavou oblast, kde se business intelligence a analytika rychle rozvíjejí?

**PP:** Je to například oblast lidských zdrojů. Zde jsou velmi aktuální tzv. retenční modely, které se snaží omezit riziko, že daný zaměstnanec odejde.

### Znamená to, že vaše řešení navrhne v pravý okamžik zvýšení mzdy, aby se předešlo výpovědi?

**PP:** Mnohem více. Mzda je jen jedním z mnoha faktorů, ovlivňujících retenci. Model dokáže nejen indikovat riziko odchodu zaměstnance, ale umí označit ty faktory, které mají na retenci u konkrétního zaměstnance největší vliv. Na základě toho můžeme navrhnout relo-



**Pavel Pilař,**  
Managing Consultant  
a Team Leader R&A,  
Partner Sophia Solutions

Je absolventem Elektrotechnické fakulty Českého vysokého učení technického v Praze. Oblasti business intelligence věnuje už více než 18 let. Jeho specializací jsou reportingové a analytické systémy. Dlouhodobě se zaměřuje na oblast dataminingu a statistické analýzy dat. V současnosti ve společnosti Sophia Solutions vede tým, který se touto problematikou zabývá. Od roku 2015 je jedním z partnerů firmy.

kaci na jinou pobočku, pokud by se jako riziko ukázala dojezdová vzdálenost. Můžeme navrhnout povýšení, vyslání na zahraniční služební cesty, nebo naopak jejich omezení. Můžeme navrhnout přeřazení do jiného oddělení, pokud se ukáže, že vztahy na pracovišti jsou tím klíčovým vlivovým faktorem.

### Musejí zaměstnanci vyplňovat nějaké dotazníky?

**JB:** Ne, tyto modely jsou založeny na datech, které firma již někde má, jen jsou opět nezkonsolidovaná na různých mís-

tech. V úvahu se berou údaje z HR modulu firemního informačního systému, z evidence docházky, hodnocení vedoucího atd. Jako faktor se může ukázat změna v intenzitě e-mailové korespondence nebo změna chování na internetu. Pro sestavení modelu je potřeba mít data za větší skupinu zaměstnanců na srovnatelné pozici – třeba v koncovém prodeji nebo na výrobní lince.

### Nebojí se zákazníci, že si koupí drahé BI řešení, které jim pak bude k ničemu?

**PP:** Především u prediktivních modelů děláme stále častěji tzv. proof of concept, tedy něco jako ověřovací řešení. Při jeho tvorbě použijeme jen klíčovou část dat a menší počet prediktorů. Výsledný model má pochopitelně nižší spolehlivost oproti plnému modelu. Ale i z něj se dá spolehlivě poznat, že taková cesta má smysl. Důležité je, že zákazníci dostanou do ruky nástroj, který si mohou prakticky vyzkoušet a ověřit jeho použitelnost. Základní výhodou jsou náklady na proof of concept, které často nedosahují ani desetin nákladů na plné řešení. Navíc vynaložené prostředky nejsou ztraceny, většina úkonů nutných pro ověřovací model se využije pro finální řešení.

### Je něco, co vás odlišuje od konkurence?

**JB:** Soustředujeme se jen na data, jejich integraci, konsolidaci a následně jejich analýzu. Nejsme závislí na jedné nebo dvou technologiích; podporujeme platformy většiny velkých vendorů. Je to náročnější cesta, ale jednoznačně z toho může profitovat zákazník. Můžeme mu tak doporučit řešení, jež je nevhodnější pro splnění konkrétních požadavků nebo jež zapadá do jeho vlastní infrastruktury, kterou má k dispozici.

**PP:** To, co děláme, děláme rádi a baví nás to. Jsem přesvědčen, že se to pozitivně odráží i na vztahu s našimi zákazníky. Určitě to je i jeden z důvodů, proč řada z nich s námi spolupracuje dlouhodobě. ■

ALEXANDER LICHÝ

**Jsme schopni propojit data z různých systémů tak, abychom dali managementu mnohem širší pohled na celek. Výsledkem může být výrazná konkurenční výhoda.**