## 13. Big Data

*Nástup webu, mobilních zařízení a dalších technologií zapříčinil zásadní změnu charakteru dat a způsobu jejich využití. Mnohé společnosti si dnes uvědomují, jak důležitou roli mohou pro dosažení jejich obchodních cílů hrát právě Big Data. Může jít např. o identifikaci faktorů, které ovlivňují rozhodnutí zákazníků o nákupu. Informace získané analýzou Big Data tak mohou firmám pomoci zlepšit jejich fungování, zajistit lepší konkurenceschopnost a růst jejich výnosů.*

*Pozn.: otázky v této kapitole se vztahují k roku předcházejícímu šetření, tj. zde konkrétně roku 2019*

## Hlavní zjištění

* Analýze Big Data, tedy dat obrovského obsahu vznikajících prakticky nepřetržitě z nejrůznějších elektronických činností nebo záznamů, se v roce 2019 věnovala necelá **desetina firem v ČR**. Big Data analyzují především **velké podniky s více než 250 zaměstnanci** (25 %) a z hlediska odvětví firmy působící v oboru nazvaném Tvorba programů a vysílání(CZ-NACE 60), což znamená přenos dat, která jsou součástí rozhlasového a televizního vysílání, zpravodajství apod. nebo firmy z odvětví doprava a skladování, nejčastěji vyzvedávání, doprava a doručování zásilek (CZ-NACE 53). V každém z těchto dvou odvětví analyzovalo v roce 2019 Big Data více než 30 % subjektů.
* Big Data se dají pro analýzu a další využití získávat z nejrůznějších zdrojů. Nejvíce firem v ČR analyzuje **data o poloze** pocházející např. z přístrojů využívajících GPS. Tato data analyzovalo v roce 2019 celkem 5 % firem s 10 a více zaměstnanci, tj. více než polovina firem, které analyzovaly v roce 2019 Big Data. Zřejmě nepřekvapí, že nejčastěji data o poloze využívají firmy z odvětví doprava a skladování a z něj vyzvedávání, doprava a doručování zásilek (CZ-NACE 53), na druhém místě subjekty z odvětví Pozemní a potrubní doprava (CZ-NACE 49).
* Druhými nejčastěji využívanými zdroji Big Data jsou **data z jiných zdrojů** (např. data ze **zákaznických databází**, data z internetových vyhledávačů, které slouží k **získání informací o chování uživatelů** internetu – např. analýza používaných klíčových slov nebo jiná **otevřená data veřejně dostupná na webu**). Tento typ velkých dat analyzovala v roce 2019 celkem 4 % firem, tedy bezmála polovina subjektů, které analyzovaly Big Data. Nejčastěji šlo o subjekty z odvětvové sekce CZ-NACE J Informační a komunikační činnosti, především pak Tvorba programů a vysílání (CZ-NACE 60) nebo Činnosti v oblasti IT (CZ-NACE 62-63). Analyzování tohoto typu velkých dat se věnovala více než pětina firem z odvětví Reklama a průzkum trhu (CZ-NACE 73).
* Třetí nejčastěji využívaný zdroj Big Data byl v roce 2019 prostor **sociálních sítí a jiných sociálních médií**. Získávání informací o uživatelích těchto nových médií a vytěžování obchodních poznatků, které uživatelé sociálních médií sdílí, se věnovala více než 3 % všech firem, tedy více než třetina firem analyzujících Big Data. Tento zdroj velkých dat využívaly nejčastěji subjekty zabývající se rozhlasovým a televizním vysíláním, zpravodajstvím, tedy subjekty, které se sociálními médii velmi intenzivně pracují.
* **Data o výrobních a logistických činnostech z „chytrých“ senzorů** (např. údaje o aktuální spotřebě energie, materiálu, opotřebení nebo pohybu zboží), byly v roce 2019 nejméně častým zdrojem Big Data (2 % všech firem, tedy přibližně čtvrtina firem analyzujících Big Data). S těmito zdroji velkých dat pracovaly v roce 2019 nejčastěji subjekty zabývající se dopravou a doručováním zásilek (CZ-NACE 53) a také firmy z automobilového průmyslu (CZ-NACE 29).
* Pokud firma analyzuje Big Data, ukazuje se, že se této činnosti častěji věnují **vlastní zaměstnanci firmy** (případně zaměstnanci mateřské firmy) **než najatí externí specialisté**. Vlastním zaměstnancům svěřují analýzu velkých dat především velké nebo středně velké firmy. Vlastní zaměstnance upřednostňují pro analyzování velkých dat nejčastěji firmy z IT odvětví nebo z mediálního sektoru. Na zaměstnance externího subjektu se častěji než na vlastní zaměstnance obracejí firmy působící ve stravování a pohostinství, v ubytování a firmy působící v obchodu (CZ-NACE 45).
* V šetření probíhajícím v roce 2020 byly do modulu Big Data zařazeny nové otázky týkající se **metod pro analýzu Big Data**. Otázky zjišťovaly využívání metody strojového nebo hlubokého učení, dále využívání počítačového zpracování, generování či rozpoznávání lidské řeči. Tyto metody nejsou zatím v českém podnikatelském prostředí příliš využívány. **Strojové učení** použila v roce 2019 necelá dvě procenta firem, nejčastěji subjekty z odvětví Činnosti v oblasti IT (CZ-NACE 62) a **metodu počítačového zpracování přirozeného jazyka**, neboli analýzám či generování textů či mluveného slova 2,6 % firem, nejčastěji firmy z odvětví Reklama a průzkum trhu (CZ-NACE 73).
* Nejčastěji využívanou metodou analýzy Big Data byly v roce 2019 **jiné metody, např. analýza obrazových informací, prediktivní analýzy nebo nejrůznější analýzy webu**. Tyto metody využilo v roce 2019 více než 6 % firem a přibližně pětina velkých subjektů. Z odvětvového hlediska využívají jiné metody nejčastěji firmy zabývající se reklamou a průzkumy trhu, tvorbě programů a vysílání nebo činnostmi v oblasti IT.
* České firmy byly **v mezinárodním srovnání[[1]](#footnote-1)** vevyužívání Big Data za rok 2017 **pod průměrem EU28**. Evropskému žebříčku dominovaly Malta, Nizozemsko, Belgie, Irsko a Finsko. V každé této zemi se analýze Big Data věnovala více než pětina subjektů. Ukazuje se také, že v jiných evropských zemích firmy častěji než v ČR analyzují Big Data pocházející ze sociálních sítí a jiných sociálních médií i data o poloze z přenosných zařízení.

**Tab. 13.1: Firmy s 10 a více zaměstnanci v ČR analyzující Big Data v roce 2019**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Firmy analyzující Big Data celkem | zdroje Big Data | Kdo pro firmy analyzoval Big Data |
| data o výrobních/ logistických činnostech z "chytrých" senzorů  | data o poloze z přenosných zařízení (např. GPS komunikace) | data ze sociálních sítí a jiných sociálních médií | data z jiných zdrojů (např. zákaznické databáze) | zaměstnanci firmy nebo mateřské firmy  | externí zpracovatelé (např. zaměstnanci jiného subjektu) |
| **Firmy celkem (10+)** | **9,1**  | **2,1**  | **5,3**  | **3,4**  | **4,2**  | **6,7**  | **4,4**  |
| **Velikost firmy** |   |   |   |   |   |   |   |
|  10–49 zaměstnanců | 7,4  | 1,2  | 4,4  | 2,8  | 3,4  | 4,9  | 3,8  |
|  50–249 zaměstnanců | 13,5  | 3,5  | 7,8  | 4,9  | 6,3  | 11,1  | 5,2  |
|  250 a více zaměstnanců | 24,8  | 14,4  | 12,2  | 8,3  | 12,2  | 22,3  | 11,7  |
| **Odvětví (ekonomická činnost)** |   |   |   |   |   |   |   |
|  Zpracovatelský průmysl | 8,0  | 3,5  | 4,3  | 2,6  | 3,0  | 6,6  | 3,1  |
|  Výroba a rozvod energie, plynu, tepla | 12,6  | 2,5  | 7,8  | 3,9  | 2,0  | 9,1  | 4,5  |
|  Stavebnictví | 6,4  | 0,5  | 6,2  | 0,6  | 1,5  | 3,6  | 2,8  |
|  Obchod a opravy motorových vozidel | 11,8  | 2,1  | 3,2  | 6,3  | 6,6  | 5,3  | 9,1  |
|  Velkoobchod | 8,2  | 0,6  | 6,1  | 3,8  | 5,7  | 5,2  | 5,1  |
|  Maloobchod | 12,0  | 0,9  | 3,1  | 8,0  | 6,5  | 8,9  | 7,5  |
|  Doprava a skladování | 15,1  | 3,4  | 14,9  | 1,1  | 2,0  | 8,6  | 6,9  |
|  Ubytování | 8,1  | 1,3  | . | 5,7  | 4,2  | 4,6  | 5,1  |
|  Stravování a pohostinství | 2,7  | . | 0,9  | 2,0  | 2,0  | 1,4  | 2,4  |
|  Činnosti cestovních agentur a kanceláří | 11,1  | . | . | 10,2  | 5,5  | 8,5  | 8,3  |
|  Audiovizuální činnosti; vydavatelství | 16,5  | . | 4,8  | 11,8  | 13,3  | 14,0  | 8,6  |
|  Telekomunikační činnosti | 11,7  | 6,6  | 8,3  | 4,4  | 7,8  | 9,9  | 4,3  |
|  Činnosti v oblasti IT | 27,0  | 7,9  | 9,8  | 10,5  | 21,1  | 26,9  | 6,1  |
|  Činnosti v oblasti nemovitostí | 3,8  | 1,4  | 1,7  | 1,8  | 3,1  | 2,8  | 1,6  |
|  Profesní, vědecké a technické činnosti | 9,0  | 1,8  | 3,4  | 5,3  | 6,6  | 8,3  | 4,6  |
|  Ostatní administrativní a podpůrné činnosti | 7,4  | 0,8  | 5,9  | 1,8  | 2,3  | 4,3  | 4,0  |

*Pozn. tečky v tabulce značí nespolehlivé údaje (velmi malý počet firem v daných velikostních a odvětvových skupinách)*

*podíl na celkovém počtu firem s 10 a více zaměstnanci v dané velikostní a odvětvové skupině (v %)*

Zdroj: Český statistický úřad, 2020

**Graf 13.1:** **Firmy s 10 a více zaměstnanci v ČR analyzující Big Data**

*podíl na celkovém počtu firem s 10 a více zaměstnanci v dané velikostní skupině v daných letech (v %)*

**Graf 13.2:** **Zdroje Big Data využívané firmami s 10 a více zaměstnanci v ČR v roce 2019**

*podíl na celkovém počtu firem s 10 a více zaměstnanci v dané velikostní skupině v daných letech (v %)*

Zdroj: Český statistický úřad, 2020

**Graf 13.3:** **Firmy s 10 a více zaměstnanci v zemích EU analyzující Big Data v roce 2017**

*podíl na celkovém počtu firem s 10 a více zaměstnanci v dané zemi z*droj dat: Eurostat, leden 2020

**Graf 13.3:** **Využívané metody analýzy Big Data ve firmách v s 10 a více zaměstnanci v ČR v roce 2019**

*podíl na celkovém počtu firem s 10 a více zaměstnanci v dané velikostní skupině v daných letech (v %)*

**Graf 13.4: Kdo pro firmy s 10 a více zaměstnanci v ČR analyzoval Big Data v roce 2019**

*podíl na celkovém počtu firem s 10 a více zaměstnanci v dané velikostní a odvětvové skupině (v %)*

Zdroj: Český statistický úřad, 2020

1. Zdrojem dat pro mezinárodní srovnání je databáze Eurostatu, která byla naposledy aktualizována v lednu 2020. Poslední dostupné mezinárodní údaje k 3D tisku se ale vztahují k roku 2017: [http://ec.europa.eu/eurostat/web/digital-economy-and-society/data/comprehensive-database](http://ec.europa.eu/eurostat/web/digital-economy-and-society/data/comprehensive-database%20%20)  [↑](#footnote-ref-1)