

**BIG DATA PRE PODPORU  
ROZHODOVACÍCH PROCESOV PRI  
RIEŠENÍ ZAŤAŽENIA ČLOVEKA  
RIZIKOVÝMI FAKTORMI PROSTREDIA**

Ing. Daniela Borošová, PhD.

Ing. Daniel Mikláš

# ČO SÚ BIG DATA

**Big dáta** - pojem pre súbory dát, ktorých rozsah je natoľko veľký, že je ťažké ich spracovávať v rozumnom čase tradičnými databázovými nástrojmi alebo aplikáciami.

Typické pre big dáta - 3V:

- **Veľkosť** či **objem** (**V**olume) - veľkosť desiatok terabajtov (TB) až po petabyty (PB).
- **Veľká rozmanitosť** (**V**ariety) - rôzne typy dát, najmä neštruktúrovaných dát, metadát, geodát
- **Veľká rýchlosť** (**V**elocity) akou dáta vznikajú

vznikajú pri náročných vedeckých výpočtoch a meraniach a tiež v dôsledku **erupcie digitálneho obsahu** v ostatných rokoch

- zvýšené využívanie webových stránok, aplikácií, sociálnych sietí, mobilných zariadení a ďalších technológií,
- prehliadanie webových stránok, komunikácia na sociálnych sieťach alebo v diskusných fórach, nákupy na webshop, platby platobnou, pri takmer akomkoľvek využívaní inteligentných telefónov, vytváranie a zdieľanie multimedialného obsahu ako sú fotky, videá či hudba.
- Tento trend bude v budúcnosti ešte viac urýchlený nástupom internetu vecí.

<https://managementmania.com/sk/big-data>

# BIG DÁTA V PRAXI

- **dopĺňajú tradičné zdroje dát (pochádzajúce z ERP, CRM či ďalšie podnikových systémov) a rozširujú spektrum dát pre rozhodovanie a riadenie podnikov a organizácií.**
- **dáta o zákazníkovi alebo zákazníckych segmentoch, či pomôcť nájsť ich vzorce správania.**
- **veľké množstvo užitočných dát, ktoré sú v “tradičných” podnikových aplikáciách. Spracovanie a analýzy veľkých dát nachádzajú uplatnenie najmä v oblastiach s rastúcim objemom dát z prevádzkových systémov.**
- **Typickým príkladom sú obchodné reťazce (dáta o správaní zákazníkov, o ich nákupných zvyklostiach), bankovníctva, mobilní operátori, prevádzkovatelia webových stránok, on-line služieb, webových aplikácií alebo napríklad oblasť finančných a kapitálových trhov.**
- **big data - v oblastiach vedy, výskumu alebo napríklad zdravotníctvo, kde vniká obrovské množstvo dát z rôznych meracích prístrojov a kde možno analyzovať aktuálne aj historické dáta a vytvárať súvislosti alebo ich využívať pri simulácii.**

# BIG DATA VO VEREJNOM ZDRAVOTNÍCTVE

**rozsiahle súbory elektronických údajov,**

- ktoré sú heterogénne,
- multimodálne,
- štruktúrované,
- neštruktúrované,
- alebo neúplné.

**Verejné zdravotníctvo produkuje veľké objemy údajov a rovnako získava, integruje a spracováva rozsiahle súbory údajov aj z iných rezortov.**

- v blízkej budúcnosti sa očakáva výrazný nárast týchto údajov.
- analýza jednej ľudskej vzorky chromatografickou analýzou s vysoko rozlišovacím hmotnostným spektrometrom vyprodukuje z jedného merania 50 až 150 MB údajov. Na plnú charakteristiku vzorky sa vyžaduje 4 až 10 takýchto meraní.

# BIG DATA VO VEREJNOM ZDRAVOTNÍCTVE

## **rôznorodosť**

- sa označujú ako vnútorné (osobnostné) rizikové faktory
- vonkajšie rizikové faktory.

## **obsahujú napríklad:**

- klinické údaje obyvateľov,
- genetické údaje,
- údaje z laboratórnych vyšetrení,
- údaje o životnom štýle, vzdelaní, a zdravotnej starostlivosti,
- údaje zo životného prostredia (o jednotlivých zložkách a faktoroch životného prostredia ako sú ovzdušie, voda, prírodná a umelá radiácia,...),
- údaje z pracovného prostredia,
- socioekonomické údaje.

# BIG DATA VO VEREJNOM ZDRAVOTNÍCTVE

K týmto údajom **nie je možné** pristupovať tradičnými metódami alebo nástrojmi na spracovanie údajov.

## Vývoj programového vybavenia a technológií

- návrh architektúry a vývoj modelu Registra humánneho biomonitoringu
- modelu Digitálneho archívu vzoriek prostredia,
- prepojenie týchto systémov s laboratórnymi systémami, (vrátane chromatografickej analýzy s vysoko rozlišovacou hmotnostnou spektrometriou), so znalostnou databázou,
- výskum analytických metód na hľadanie odpovedí o možných príčinných súvislostiach medzi expozíciou negatívnym faktorom a vznikom ochorení,
- výskum prediktívnych štatistických modelov pravdepodobnosti vzniku ochorení,
- výskum modelov využívajúcich pravdepodobnostný a matematicko-štatistický aparát,
- výskum grafickej a geografickej vizualizácie dát výsledkov projektu.

# VÝZVA O NENÁVRATNÝ FINANČNÝ PRÍSPEVOK

**Finančné prostriedky z Európskej únie je možné čerpať v rámci operačného programu Výskum a inovácie**

**Skratka OPVAI**

[www.opvai.sk](http://www.opvai.sk)



**Výzvy na projekty operačného programu Výskum a inovácie vypísala Výskumná agentúra Ministerstva školstva**

[www.vyskumnaagentura.sk](http://www.vyskumnaagentura.sk)



# VÝZVA O NENÁVRATNÝ FINANČNÝ PRÍSPEVOK

**Konzorcium partnerov má vybudovanú vývojovú základňu s dostatočným výpočtovým výkonom pre spracovanie veľkých objemov údajov, na analýzu big data a na vývoj dátových prototypov. Konzorcium má vybudovanú aj laboratórnu základňu na analýzu ľudských vzoriek a vzoriek prostredia.**

z podnikateľského prostredia

- Softec, spol. s r. o.,
- Asseco Central Europe a.s.,
- Environmental Institute, s. r. o.,

z verejnej správy

- Regionálny úrad verejného zdravotníctva v Banskej Bystrici,
- Úrad verejného zdravotníctva Slovenskej republiky

z vedeckého prostredia

- UVP UNIZA, Žilina
- TU Zvolen,
- JLF UK Martin.



# VÝZVA O NENÁVRATNÝ FINANČNÝ PRÍSPEVOK

**Projektový klaster je založený na dlhodobej spolupráci, tvorbe technologických platforiem, so silnými väzbami medzi spoločnosťami a inštitúciami, pričom všetci účastníci hrajú rovnako dôležitú úlohu.**

- Klaster produkuje znalosti, ktoré si v projektovom tíme odovzdáva a vymieňa, a tým podporuje možnosti rastu projektového tímu založenom na zvyšovaní inovačnej schopnosti a excelentnosti vo výskume a inováciách so zámerom podporiť udržateľný rast zamestnanosti a kvality života. Projektový klaster podporí tvorbu nových pracovných výskumných miest a v budúcnosti aj tvorbu nových spoločných projektov.

# VÝZVY NA PODPORU DLHODOBÉHO STRATEGICKÉHO VÝSKUMU A VÝVOJA

Projektová doména	Indikatívna výška finančných prostriedkov z Európskeho fondu regionálneho rozvoja
DOPRAVNÉ PROSTRIEDKY PRE 21. STOROČIE	63 072 000 €
PRIEMYSEL PRE 21. STOROČIE	107 712 000 €
ZDRAVÉ POTRAVINY A ŽIVOTNÉ PROSTREDIE	36 000 000 €
<b>DIGITÁLNE SLOVENSKO A KREATÍVNY PRIEMYSEL</b>	<b>47 520 000 €</b>
ZDRAVIE OBYVATEĽSTVA A ZDRAVOTNÍCKE TECHNOLOGIE	33 696 000 €
SPOLU	288 000 000 €

# OBSAHOVÉ ZAMERANIE PROJEKTU

Hlavný trend	Vedľajší trend	Produktová línia
II. Digitálne technológie pre spoločnosť	2. Systémové nástroje digitalizácie	<ul style="list-style-type: none"><li>• Vývoj programového vybavenia a technológií pre inteligentné výrobné systémy, komponenty a uzly, ako aj mestá a komunity;</li><li>• <b>Služby a riešenia v oblasti spracovania veľkých objemov dát, rýchle spracovanie dát (Big Data, High performance computing, cloud computing edge computing);</b></li><li>• Uchovávanie a sprístupňovania informácií (Open Data, Linked data);</li></ul>

# VÝZVA O NENÁVRATNÝ FINANČNÝ PRÍSPEVOK



EURÓPSKA ÚNIA  
Európsky fond regionálneho rozvoja



Európsky fond  
regionálneho rozvoja.

spolufinancovanie zo  
strany  
prijímateľ'a/partnera



MINISTERSTVO  
FINANCIÍ  
SLOVENSKEJ REPUBLIKY

štátny rozpočet

# INTENZITA PRÍSPEVKU

V PRÍPADE ŽIADATEĽOV A PARTNEROV, KTORÍ SÚ VÝSKUMNÝMI INŠTITÚCIAMI SA V RÁMCI PROJEKTU BUDE REALIZOVAŤ NEZÁVISLÝ VÝSKUM A VÝVOJ

Žiadateľ/partner (výskumná inštitúcia)	Maximálna intenzita príspevku v (%)	Zdroje financovania v pomere EFRR / štátny rozpočet SR	Minimálna výška spolufinancovania zo zdrojov Žiadateľa/partnera v (%)
<b>Organizácie štátnej správy</b>	<b>100 %</b>	85 % / 15 %	<b>0 %</b>
Ostatné subjekty verejnej správy	95 %	85 % / 10 %	5 %
Mimovládne / Neziskové organizácie	95 %	85 % / 10 %	5 %
Súkromné vysoké školy vrátane právnických osôb vykonávajúcich výskum a vývoj zriadených súkromnými vysokými školami	90 %	85 % / 5 %	10 %

# PROJEKT SA BUDE POSUDZOVAŤ Z POHĽADU



OP výskum a inovácie

Programové obdobie 2014 - 2020

- **Excelentnosti obsahového zámeru**
- **Miery zapojenia univerzitných vedeckých parkov**
- **Minimalizácii podielu kapitálových výdavkov na celkových výdavkoch**

# BIG DATA VO VEREJNOM ZDRAVOTNÍCTVE

**Na naplnenie výstupov projektu sú nutné odborné, vedecké a technické činnosti z rôznych oblastí:**

- **Činnosti z oblasti informačných a komunikačných technológií**
- **Činnosti z oblasti laboratórných systémov**
- **Činnosti z oblasti translačného medicínskeho výskumu**

# BIG DATA VO VEREJNOM ZDRAVOTNÍCTVE

## Činnosti z oblasti informačných a komunikačných technológií

- analýza požiadaviek,
- analýza dátových zdrojov,
- analýza a transformácia big data,
- vývoj databázy, vývoj HW a SW architektúry riešenia,
- doménová analýza,
- analýza a vývoj predikčných procesov,
- analýza a vývoj 2D a 3D vizualizácie,
- vývoj architektúry dátových prototypov, vývoj dátových prototypov,
- vývoj databázových procedúr,
- bezpečnostná analýza,
- manažovanie kontroly kvality,
- analýza testovacích procedúr , vývoj testovacích procedúr,
- testovanie, zavedenie dátových prototypov do overovacej prevádzky,
- riadenie projektových prác.



# BIG DATA VO VEREJNOM ZDRAVOTNÍCTVE

## Činnosti z oblasti laboratórných systémov

- literárna rešerš,
- vývoj metódy,
- kalibrácie,
- overovanie postupov,
- validácia postupov,
- riadenie kvality,
- proces akreditácie metód,
- zapojenie do medzilaboratórných testovaní,
- dokumentovanie postupov a vypracovanie metodík,
- zber vzoriek, analýza vzoriek,
- archivácia vybraných vzoriek,
- evidencia výsledkov vzoriek a ich zaevidovanie v dátových prototypoch.

# BIG DATA VO VEREJNOM ZDRAVOTNÍCTVE

## Činnosti z oblasti translačného medicínskeho výskumu

- hodnotenie zaťaženia obyvateľstva rizikovými faktormi zo životného prostredia,
- klasifikácia expozície a rizík,
- interpretácia rizík,
- analýza kvalitatívnych a kvantitatívnych vzťahov medzi mierou vystavenia, dávkou a jej účinkom,
- vývoj pragmatických pravidiel posúdenia rizík vrátane ich hraničných hodnôt,
- návrh politik na ochranu verejného zdravia a prevenciu chorôb.

# BIG DATA VO VEREJNOM ZDRAVOTNÍCTVE

## Predmet projektu

- výskum veľkého objemu údajov zaťaženia ľudí vonkajšími faktormi prostredia
- identifikácia škodlivých chemických faktorov vstupujúcich do ľudského organizmu.
- Výskum a vývoj digitálnych prototypov umožní analyzovanie možných príčinných vzťahov medzi expozíciou chemickým faktorom a zdravotnými následkami v podobe chorôb, poškodenia zdravia alebo zmien vo fyziologických parametroch u ľudí.
- Vstupné údaje do digitálnych prototypov budú získané z laboratórnych prístrojov mapujúcich kvalitu životného prostredia,
- z laboratórnych prístrojov mapujúcich chemické látky, alebo ich metabolity v ľudskom tele v telesných tekutinách,
- z bližšieho popisu životného štýlu osoby, ku ktorej sa vzťahuje vzorka humánneho biomonitoringu,
- zo vzťahu dĺžky expozície a intenzity expozície faktormi životného prostredia sledovanej osoby, a pod.

Ďakujem za pozornosť  
borosova@vzbb.sk