

# BIG DATA: KONEC EVALUACE, JAK JI ZNÁME (?)

Marek Havrda

2014

# BUDOUCNOST EVALUACE

Konference:

Next Generation Evaluation, 11/2013

(Stanford Social Innovation Review & FSG)

BIG DATA ... JEDNO ZE 3 „game-changing“ TÉMAT

# BIG DATA

máme k dispozici dříve nebývalý a stále rostoucí  
objem informací

před rokem 2000 byla pouze  $\frac{1}{4}$  všech uložených informací v  
digitální formě, dnes méně než 2 % jsou nedigitální

nové způsoby jejich využití

nemění se „pouze“ způsob komunikace, ale

vytváří se nové poznatky, které nebyly na základě „small data“  
dostupné

# DATAFIKACE

- dříve nedigitální se digitalizuje a zanechává digitální stopu – bankovníctví, zábava (mp3, streaming), vzdělávání (blended learning, flipped cassroom)...
- každodenní využití zařízení, která vytváří další data – chytré telefony, senzory, fotoaparáty s GPS...
- dobrovolně sdílíme informace – Facebook, Twitter, Instagram...

90 % informací v dnešním světě bylo vytvořeno v posledních DVOU letech (IBM) – jsme batolata

## BIG DATA – VYUŽITÍ ... SOUKROMÝ SEKTOR

- automatická detekce pojistných podvodů
- zkušenosti spotřebitelů s novými výrobky
- automatizace tvorby cen (slevy)
- individualizované nabídky

## BIG DATA – VYUŽITÍ .... not only PREDICTING (THE IMPACT)

- dopravní zpravodajství – Google Maps
- agregace sms a sociálních médií pro reakci na přírodní katastrofy - Ushahidi
- predikce epidemie chřipky – Google (data-based model)
- požární prevence – New York City (nepřímo související data od 17 městských organizací včetně výskytu hlodavců či pozdních plateb)
- predikce kriminality – Los Angeles
- léčebné postupy: nejlepší praxe – Autonomy
- predikce nezaměstnanosti (analýza sociálních médií) - OSN
- plánování rekvalifikací – na základě analýzy nabídky volných míst

## BUDOUCNOST EVALUACE – DATA K DISPOZICI

- sociální i „tradiční“ média
- inzerce
- dotazy vyhledávání
- pohyby SIM karet + GPS data + ... Internet of Things
- finanční transakce
- data veřejného sektoru (Open Data)

# BUDOUCNOST EVALUACE - METODOLOGIE

(na základě Foreign Affairs, 2013)

- od vzorků k N=all
- od „čistých“, ale velmi parciálních dat k akceptaci „nečistoty“ v rozsáhlých datech (epidemie chřipky)
- od snahy identifikovat závislosti (kauzality) k akceptaci souvislosti (korelace)



## BOUDOUCNOST EVALUACE - BIG DATA UMOŽŇUJÍ:

hodnotit reálné experimenty (např. porovnání dvou podobných regionů): evidence-based, management by results

omezit některé neduhy tradiční metodologie: self-reported (attitude-behaviour gap) & vzorky populace (někdy dokonce self-selected – „vrátilo se nám 10% dotazníků“)

realizovat longitudinální studie

(často) nepotřebujeme vědět dopředu, jaká data budeme pro hodnocení vlastně potřebovat

## BIG DATA A BUDOUCNOST EVALUACE

odpověď na otázku CO funguje, ne PROČ to funguje => antropologické přístupy

on-line evaluace: hodnocení programů v reálném čase s cílem jejich korekce a adaptace pro maximalizaci dopadu (x současný mid-term, ex-post (ne sedmiletka na základě tři roky starých dat))

více holistický přístup (systems, design thinking) – zlepšení evidence-based (např. vzhledem k RCTs)

dostupnost .... otevírání dat – Open Data – př. data sociálního zabezpečení (anonymizovaná a při zachování privacy)

Development Evaluation & Shared Measurement (Next Generation Evaluation)

# **Děkuji za pozornost**

**[linkedin.com/in/marekhavrda](https://www.linkedin.com/in/marekhavrda)**