

# Srečanje 1: TIPOLOGIJE VIROV PODATKOV

---

- način pridobivanja:
  - primarni (log datoteke, zbrane za naš namen)
  - sekundarni (ADP)
- namen uporabe:
  - kvantitativni podatki (nacionalni etaloni)
  - kvalitativni podatki (besedilo bloga)
- stopnja razkritja:
  - mikro podatki (o individualnih enotah, npr. neto plača posameznika)
    - neanonimizirani
    - anonimizirani (deindividualizirani)
    - statistično zaščiteni
    - za javno uporabo
  - agregirani podatki (povprečna neto plača v Sloveniji po dejavnosti zaposlitve)
- obdelanost podatkov:
  - surovi podatki (vremenska postaja)
  - prečiščeni podatki (SURS)
- geografska pokritost1:
  - lokalni (občinski, regionalni)
  - državni
  - mednarodni
  - globalni, svetovni
- geografska pokritost2:
  - slovenski (SURS)
  - tuji (Statistics Netherlands)
- geografska pokritost3:
  - evropski (SURS, Eurostat)
  - neevropski (US Census)
- število vključenih držav:
  - nacionalni (SURS, ONS = Britanski statistični urad, INSEE = francoski statistični urad)
  - mednarodni (Eurostat; World Bank)
- jezik
  - enojezični (angleški, kitajski, španski, ruski, portugalski, francoski...)
  - dvo- in večjezični
- področje
  - tehnični (ELES)
  - ekonomski (DURS)
  - naravoslovni (ARSO)
  - družboslovni (Facebook)
- raven obravnave:
  - mikro podatki (gospodinjstvo, podjetje)
  - mezo podatki (občina, dejavnost)
  - makro podatki (država)
- zavezanost Zakonu o varstvu osebnih podatkov
  - osebni podatki (katerikoli podatki, ki se nanašajo na posameznika, ne glede na obliko, v kateri so izraženi)
  - neosebni, drugi podatki
- preverjenost, zanesljivost:
  - preverjeni, recenzirani podatki (učbenik)
  - nepreverjeni podatki, podatki z neznano zanesljivostjo (Wikipedia)
- verodostojnost:
  - resnični
  - lažni, zavajajoči
- preverljivost (z vidika tretje osebe):
  - objektivni (višina)
  - subjektivni (mnenje, čustvo)
- (zavestno) sodelovanje udeležениh oseb
  - s privolitvijo in sprotim nadzorom nad posredovanimi podatki (anketiranje, intervju)
  - s privolitvijo in brez sprotnega nadzora nad posredovanimi podatki (kartice zvestobe, piškotki)
  - brez privolitve (štetje prometa)

- vrsta zapisa:
  - digitalni
  - analogni (zdravniške kartoteke)
- oblika komuniciranja oz. izražanja:
  - številski
  - tekstovni
  - slikovni
  - zvočni
  - multimedia (kot kombinacija različnih oblik podatkov)
- vključenost enot:
  - populacijski (Poslovni register Slovenije)
  - vzorčni (Raba Interneta v Sloveniji)
- sprotnost zajema:
  - registrski in registrirani podatki (CRP, skenirano trgovsko blago)
  - z zamikom (Anketa o potrošnji gospodinjstev)
- časovno sledenje:
  - presečni
  - longitudinalni
- periodika zbiranja:
  - enkratni podatki
  - večkrat zbrani podatki (različne občasne ankete)
  - periodični podatki (RIC)
- aktualnost:
  - arhivski podatki (zgodovinski zapisi)
  - sodobni podatki (socialna omrežja)
- formalne implikacije:
  - statistični podatki (Anketa o delovni sili, Statistični del Poslovnega registra Slovenije)
  - administrativni podatki (CRP)
- stroški dostopa:
  - plačljivi podatki (GURS)
  - brezplačni podatki (ESS)
- oviranost dostopa:
  - prost dostop
  - dostop z registracijo, identifikacijo
- vrsta ponudnika:
  - uradni (AJPEŠ)
  - komercialni/profitni (gvin.com)
  - nekomercialni (Oqali (transformed food products))
- sektor lastnika:
  - javni (ARSO)
  - zasebni (bančni izpiski)
- dostopnost:
  - zunanji/javno dostopni (Wikipedia, izmerjena temperatura zraka)
  - notranji/zasebni (klicane mobilne številke, izmerjena telesna temperatura)
- število uporabljenih virov
  - posamezni
  - združeni
    - avtomatsko združeni (prepletene storitve=mashup)
    - preverjeno združeni (OECD)