

Žmogiškasis veiksnys: 2014 m. ir 2015 m. Korupcijos suvokimo indekso rezultatų klaidos paaškinimas - Santhosh Srinivasan, (TI) and Finn Heinrich (buvęs TI tyrimų vadovas)

Korupcijos suvokimo indeksas (KSI) yra žinomas „Transparency International“ tyrimas ir, ko gero, labiausiai cituojamas korupcijos indeksas. KSI rezultatus naudoja politikos formuotojai, verslininkai, žurnalistai, mokslininkai ir aktyvistai.

Kartu su ekspertais KSI metodologija buvo nuolat atidžiai vertinama ir tobulinama. Visgi 2014 m. ir 2015 m. įsivėlė klaidos, atsiradusios dėl rankinio duomenų suvedimo. Kaip tai nutiko?

KSI yra sudarytas iš tam tikro skaičiaus (šiuo metu – trylikos) šaltinių. Kai kuriuos duomenis TI gauna tiesiai iš šaltinių dar prieš jų pavišimą, kiti duomenys mums tampa pasiekiami tada, kai jų autoriai juos pavišina internete.

Pirmoji KSI tyrėjo užduotis yra sukurti vieną meta-duomenų rinkinį, kuriame tilptų visi skirtingų šaltinių duomenys, susiję su korupcija tiriamose valstybėse. Jeigu dirbama su mikro lygmens duomenimis (pavyzdžiui, verslo atstovų apklausomis), reikia šiuos rezultatus agreguoti į makro, t.y. valstybės lygmens duomenis. Pradiniai duomenys būna įvairiausių formatų: nuo Excel failų ir Word dokumentų iki .jpg formato paveikslėlių, kuriuose esančią informaciją į duomenų rinkinį veda ne kompiuteris, o žmogus.

Klaidos atsirado būtent šiame žingsnyje, kai buvo kuriamas bendras duomenų rinkinys. 2015 m. klaida atsirado tyrėjui ranka kopijuojant ir į KSI duomenų rinkinį įklijuojant rezultatus iš „Bertelsmann Sustainable Governance Indicators“ (SGI) tyrimo. Tai nutiko naudojant tik valstybių pavadinimus ir balus, bet nesinaudojant duomenų apjungimo metodais, kurie remiasi vienareikšmiais valstybių atpažinimo rodikliais, kaip, pavyzdžiui, ISO kodai.

Meta-duomenų rinkinyje buvo naudojamas terminas „Korėja, Pietų“, tuo metu originaliame „Bertelsmann SGI“ faile naudotas terminas buvo „Pietų Korėja“. Įvedant duomenis į KSI duomenų rinkinį šis skirtumas nebuvo pastebėtas, todėl šalims, esančioms tarp „Korėja, Pietų“ ir „Pietų Korėja“, buvo priskirti neteisingi „Bertelsmann SGI“ balai. Dėl to visi balai nuo Korėjos (Pietų) iki Slovėnijos „pasislinko“ viena pozicija (rezultatai lentelėje). Taip pakito 11 valstybių KSI balai.

	Neteisinga			Teisinga
Pietų Korėja	7			
Latvija	6		Latvija	7
Lietuva	8		Lietuva	6
Liuksemburgas	5		Liuksemburgas	8
Malta	3		Malta	5
Meksika	7		Meksika	3
Nyderlandai	10		Nyderlandai	7
Naujoji Zelandija	8		Naujoji Zelandija	10
Norvegija	7		Norvegija	8
Lenkija	6		Lenkija	7
Portugalija	5		Portugalija	6
Rumunija	5		Rumunija	5
Slovakija	5		Slovakija	5

Slovėnija	5		Slovėnija	5
			Pietų Korėja	5

2014 metais klaida atsirado duomenis perkeliant iš „Economist Intelligence Unit“ (EIU) tyrimo, o tai iškreipė dviejų valstybių rezultatus. Kadangi šios valstybės nebuvo vertintos 2014 metų EIU tyrime, į meta-duomenų rinkinį buvo įkopijuotos 2013 metų vertės. Panaši kopijavimo-įklivavimo klaida atsirado ir 2015 metais, kai duomenys buvo įkopijuoti ne į tuos langelius.

	Neteisinga		Teisinga
Sent Lusija			Sent Lusija
Sent Vinsentas ir Grenadinai	x		Sent Vinsentas ir Grenadinai
Samoa	x		Samoa

Tai pakeitė trijų valstybių indeksų vertes. Sent Vinsento ir Grenadinų atveju balas buvo pakeistas iš 67 į 62. Samoa išėmėme iš KSI, nes ji neturi trijų nepriklausomų duomenų šaltinių, o Sen Lusiją įtraukėme į vertinimą, nes ji turi 3 duomenų šaltinius.

Suprasdami, kad skaičiuojant KSI rezultatus naudojamas rankų darbas, TI visada taikė „keturių akių“ tikrinimo principą. Vis dėlto tokio patikrinimo metu 2014 m. ir 2015 m. klaidos nebuvo pastebėtos, nes tikrinant nebuvo peržiūrimi pirminiai duomenys (buvo tik perskaiciuojami rezultatai galutiniame duomenų rinkinyje).

Suprasdami ligšiolinio tikrinimo trūkumus, nuo 2016 m. sustiprinome KSI kokybės kontrolę:

Visų pirma, peržiūrėjome visus vidinius procesus, kad sumažintume tokių klaidų atsiradimo tikimybę. Nuo šiol du nepriklausomi tyrėjai atskirai atliks KSI skaičiavimus. Be to, padidinome kompiuterizacijos ir automatizacijos lygį dirbant su neapdorotais KSI duomenimis ir kuriant meta-duomenų rinkinį.

Stiprinant išorinę KSI verifikaciją nusprendėme, kad išoriniai vertintojai ne tik peržiūrės KSI skaičiavimus, atliktus TI-S tyrėjų komandos (kaip daryta ankstesniais metais), bet ir patys apskaičiuos KSI vertes. Po šių skaičiavimų jų gauti rezultatai bus lyginami su TI-S komandos gautais rezultatais taip siekiant užtikrinti, kad balai ir reitingas visiškai sutampa.

Kaip ir visų kitų mūsų įrankių ir produkcijos atveju, mes, TI, esame įsipareigoję nuolat peržiūrėti ir atnaujinti tyrimų metodus ir metodologijas. Atlikdami KSI metodologijos peržiūrą nusprendėme, kad pristatant KSI tyrimą reikia sukurti detalesnę instrukciją, nurodančią, kaip elgtis su dešimtosiomis skaičių dalimis, ir kada tyrėjas privalo skaičių suapvalinti iki numatytos vietos po kablelio.

Visada skyrėme daug laiko ir pastangų tobulinant ir gerinant KSI metodologiją. Siekiam KSI palyginamumo ir naudojimosi paprastumo. Ateityje ir toliau stengsimės išsaugoti KSI metodologijos ir duomenų rinkimo integralumą bei prisitaikyti prie vis tobulėjančių tyrimų metodų.

Jeigu turite klausimų, prašome susisiekti su Santhosh Srinivasan (ssrinivasan@transparency.org) arba Coralie Pring (cpring@transparency.org)

Originalų įrašą anglų kalba galite rasti [čia](#).