



Користење на обемни бази на податоци како алатки во борбата против корупцијата

Автор: Герман Филков, Центар за граѓански комуникации

ноември, 2019 година

Овој документ е подготвен од страна на Платформата на граѓански организации за борба против корупција со поддршка на американскиот народ преку Агенцијата на САД за меѓународен развој (УСАИД). Мислењата изразени во овој извештај им припаѓаат на авторите и не ги изразуваат ставовите на Агенцијата на САД за меѓународен развој или на Владата на Соединетите Американски Држави.

Извршно резиме / апстракт

Како што дигиталната технологија во светот се развива и станува нераскинлив дел од животот и работењето, така големите бази на податоци со коишто располагаат земјите (особено јавната администрација) стануваат клучно јавно добро за зајакнување на антикорупциските напори. Од една страна станува збор за проактивно објавување во отворен формат на навремени, точни и квалитетни податоци, а од друга страна тоа значи користење, обработка и анализа на тие податоци во насока на зголемување на интегритетот и отчетноста на јавниот сектор.

Анализата на обемните бази на податоци може да овозможи да се дојде до такви информации до којшто човекот без примена на оваа алатка едноставно не може никогаш да дојде.

Овој документ за јавни политики го претставува концептот на користење на обемните бази на податоци во антикорупцијата, начинот на негова примена, добрите страни и предизвиците, со посебен осврт на состојбата во Република Северна Македонија, почетните чекори, примената, како и препораки за забрзување на користењето на оваа алатка во борбата против корупцијата.

Непознавањето на потенцијалот и можната огромна улога на обемните бази на податоци во антикорупциските напори доведува до мали инвестиции во оваа област и неземање предвид на оваа улога во антикорупциските стратегии.

За да можат обемните бази на податоци со сиот свој потенцијал да се користат во антикорупциски цели, потребно е зајакнување на човечките капацитети во институциите што ја извршуваат оваа работа. Овде пред сè се мисли на обезбедување доволно техничко и информатичко знаење и опрема.

Кај нас, пак, се чини дека е потребна и позначајна политичка воља, особено на прагот на примената на новата Национална стратегија за спречување на корупцијата и судирот на интереси 2020-2024.

Содржина

Вовед	4
Нова алатка во светот за борба против корупцијата – анализа на обемните бази на податоци	5
Можности за примена на обемните бази на податоци како алатки за спречување на корупцијата кај нас	8
Заклучоци и препораки	10
Профил на авторот	10
Библиографија/користена литература	11

Вовед

Со секој изминат ден во последните години, луѓето стануваат сè позависни од дигиталната технологија. Таа не само што влијае врз успехот на владите и на бизнисите, туку секојдневно создава и нови можности за унапредување на животот и работењето.

Од друга страна, големите бази на податоци со коишто располагаат земјите (особено јавната администрација) стануваат клучно јавно добро за зајакнување на антикорупциските напори преку овозможување зголемена транспарентност, доверба во властите, интегритет и отчетност, динамизирање на граѓанското учество и соработката меѓу вклучените страни.

Притоа, концептот за „користење на обемните бази на податоци во антикорупцијата“ се поставува пошироко. Од една страна, тој опфаќа проактивно објавување во отворен формат на навремени, точни и квалитетни податоци, а од друга страна значи користење, обработка и анализа на тие податоци во насока на зголемување на интегритетот и отчетноста на јавниот сектор.

Иако и претходно имаше обиди за докажување на значењето на користењето на обемните бази на податоци во антикорупциски цели (особено преку иницијативата за Отворено владино партнерство), сепак вниманието на антикорупциската јавност во светот кон оваа проблематика особено го сврти групата на 20-те најголеми економии во светот (Г-20) која во 2017 година не само што ја прогласи борбата против корупцијата како највисок приоритет, туку се заложи за користење на иновативни решенија и нови технологии во таа борба.

Станува збор за следниве шест принципи што се однесуваат на отворените податоци и борбата против корупцијата:

- (1) отвореност по правило (пристапот до јавни информации е широко прифатен како алатка за зголемување на транспарентноста и како антикорупциска алатка, но отворените податоци одат чекор подалеку бидејќи подразбираат податоците да се направат секому достапни без да бидат побарани);
- (2) навременост и сеопфатност (објавување стандардизирани, точни, навремени и сеопфатни податочни сетови);
- (3) достапност и употребливост (озможување лесен влез и пристап за сите со што се зголемува користењето на тие податоци);
- (4) споредливост и интероперабилност (озможување на споредба и следливост на податоците од разни сектори);

- (5) за подобро управување и за ангажирање на граѓани (граѓаните да имаат увид во работењето на владините институции, да ги разберат и да учествуваат во процесите на донесување одлуки); и
- (6) за инклузивен развој и иновации (отворените податоци, преку зголемување на транспарентноста и интегритетот можат да помогнат градење на ефективни, отчетни и одговорни институции, како и да го олеснат развојот на нови модели и иновативни дигитални стратегии).

Во следните две години, по усвојувањето на овие принципи, користењето на обемните бази на отворени податоци стана една од алатките која се смета за клучна во борбата против корупцијата и чија примена, која засега е ограничена, ќе биде одлучувачка за ефикасна антикорупциска борба.

Се смета дека обемните бази на податоци особено им даваат моќ на агенциите за борба против корупцијата, на институциите за спроведување на законот, на невладините организации и на новинарите за поефикасно извршување на нивната општествена улога во намалување на корупцијата. Како еден од најпознатите примери во светот за успешна примена на анализата на обемните бази на податоци во антикорупциски цели беше случајот со обработката и анализата на 11,5 милиони документи од т.н. „Панама пејперс“.

Нова алатка во светот за борба против корупцијата – анализа на обемните бази на податоци

Обемните бази на податоци¹ (big data) претставуваат огромни количества на податоци кои настануваат од активности што се спроведени електронски преку компјутерска комуникација. Сепак, нивната количина е толку голема што ги преминува можностите на стандардно користениот софтвер за нивно складирање, обработка и управување.

Обемните бази на податоци се карактеризираат со т.н. „3V“, според почетната буква од англиските зборови на овие три главни карактеристики, односно „обем“ (Volume), „разновидност“ (Variety) и „брзина“ (Velocity). Притоа, под обем се подразбира масивноста на овие податоци, под разновидност тоа што тие можат да бидат во различен формат и структура (неструктурирани, полуструктурирани и структурирани), додека брзината се однесува на големата брзина со која овие податоци се создаваат, стануваат достапни и се менуваат со текот на времето.

¹ Забелешка на авторот: Конечниот избор на терминот „обемни бази на податоци“ е на авторот, во консултација со Катерина Димишкова, преведувачка од англиски јазик и Татјана Б. Ефтимоска, лекторка по македонски јазик.

Иако паралелно со идеите за користење на овие податоци и со нивното практично користење, всушност постојат и низа прашања од законска, етичка и техничка природа, сепак, неспорен е придонесот што анализата на обемните бази на податоци го има и може да го има во подобрување на економскиот развој, животниот стандард и – спречувањето на корупцијата, измамите и злоупотребите.

Се смета дека со појавата на обемните бази на податоци и воопшто во денешната ера на интернетот, всушност, се појавуваат и нови облици и начини за спречување на корупцијата коишто досега биле непознати и недостапни. Во најмала рака, тие можат да се користат макар за потврда или за докажување на претходно (и со други средства) откриено коруптивно однесување.

Непознавањето на потенцијалот и можната огромна улога на обемните бази на податоци во антикорупциските напори доведува до мали инвестиции во оваа област и неземање предвид на оваа улога во антикорупциските стратегии.

Анализата на обемните бази на податоци може да овозможи да се дојде до такви информации до којшто човекот без примена на оваа алатка едноставно не може никогаш да дојде.

Една од најголемите предности на обемните бази на податоци е што овозможуваат вкрстување, споредување и анализа на податоците од разни извори.

Исто така, анализата на обемните бази на податоци може да покаже случаи на отстапувања од долга низа на редовни појави и норми коишто можат да бидат случаи на корупција. Во оваа насока можат да бидат и евентуални слични шеми на однесувања што се повторуваат кај различни, неповрзани субјекти. Ова особено може да се примени во анализа на обемни бази на податоци од даночната сфера. Но, сево ова не би можело да се забележи и да се открие при редовни и обични анализи на податоци. Како област во која е особено потребно да примени анализата на обемните бази на податоци се посочува државната ревизија, која во некои земји не може да се замисли без ваква анализа.

Анализата на обемните бази на податоци може да се користи и за мерење на корупцијата. Така, експерти на Меѓународниот монетарен фонд, изработиле алгоритам за пребарување по земји на 665 милиони меѓународни новинарски текстови за мерење на корупцијата и антикорупцијата (од 30 земји, објавени во периодот од 1995 до 2017 година). Авторите на ова истражување сметаат дека веќе станува збор за т.н. трета генерација на мерење на корупцијата, со помош на користење на обемните бази на податоци, по првата генерација која ги користеше перцепциите и втората генерација која ги користеше виктимизацијата и индикаторите. Покрај овој пристап, во третата генерација

кога ги користи обемните бази на податоци се вбројуваат и некои други алатки кои веќе се користат во светот како што се мобилни апликации за пријавување на ситната корупција (Индија) и анализи во јавните набавки (Велика Британија).

Во британскиот истражувачки проект, на пример, селектирани се десетина индикатори за ризик во процесот на јавни набавки за да се анализираат половина милион тендери во повеќе од 100 земји во развој во текот на последните 20 години во рамките на помошта што овие земји ја добивале од неколку меѓународни организации. Секако, во овој конкретен проект, не можела да се докаже корупцијата, но барем биле откриени најризичните ситуации и моделите на корупција што овозможило преземање заштитни мерки и стратегии во иднина. Слични активности за спречување на корупцијата врз основа на анализа на обемна база на податоци од сите доделени проекти, применува и Европската инвестициска банка.

Во тендерите, се следи однесувањето на фирмите за да се види дали има некакви шеми во смисла на тоа: две фирми секогаш да се јавуваат на исти тендери со тоа што едната секогаш добива, а другата секогаш е втора; или три фирми да учествуваат на тендери со тоа што еднаш добива првата, друг пат втората и трет пат третата и тоа да се повторува циклично на сите тендери на коишто учествуваат; или конкурираат фирми со слични имиња зад кои можеби стојат исти или поврзани сопственици; итн.

Корејската комисија за фер трговија, пак, користи анализа на огромна база на податоци од неколку институции за да дојде до проценка на веројатноста за незаконско договарање на фирмите на пазарот и во јавните набавки.

Уште една технолошка алатка што сè почесто се споменува и се користи во спречувањето на корупцијата во оваа смисла е т.н. „blockchain“ технологија, односно блокови на информации и податоци за завршени трансакции поврзани меѓусебно во синџир без хиерархија, посредници или централен орган кој го поседува, складира или управува со него. Сите трансакции се регистрирани во шифрирана форма на компјутерите на сите учесници во системот, што обезбедува отвореност, транспарентност и непроменливост на податоците.

Токму ова овозможува од една страна спречување на корупцијата, а од друга страна дури и нејзино полесно откривање. Прво, овозможува елиминирање на посредниците во многу случаи на добивање на јавни услуги со што се елиминира и можноста за одредени форми на корупција. Натаму, имајќи предвид дека во секој од поврзаните блокови има целосна следливост на сите трансакции кои се чуваат по хронолошки ред, се оневозможува корупција со скриени и менувани трансакции. Крајно, се оневозможува и корупција поврзана со централизација и злоупотреба на моќта бидејќи станува збор за целосно децентрализиран систем на информации.

Сепак, и покрај бројните предности од користењето и анализата на обемните бази на податоци, има бројни недостатоци и пречки што се исправаат на патот на почестото, полесното и помасовното користење.

За да можат базите на податоци воопшто да се користат за каква и да е анализа, тие мора да бидат прибрани на стандардизиран начин, да бидат точни и да постои утврдена рамка за нивно споделување. Исто така, неопходно е сите овие податоци да бидат во отворен формат.

Отворените податоци играат клучна улога во транспарентна и отчетна размена на информации меѓу властите и граѓаните, но исто така и во намалување на злоупотребите и корупцијата. Меѓутоа, важни се неколку аспекти, односно овие податоци да бидат квалитетни, достапни и да можат да се разменуваат. Отворените податоци им овозможуваат на граѓаните, граѓанските организации, новинарите и истражувачите увид во владините активности, одлуки, постапки и трошења што, пак, овозможува подобар мониторинг од една страна, и поотчетни институции од друга.

За да можат обемните бази на податоци со сиот свој потенцијал да се користат во антикорупциски цели, исто така, потребно е зајакнување на човечките капацитети во институциите што ја извршуваат оваа работа. Овде пред сè се мисли на обезбедување доволно техничко и информатичко знаење и опрема.

Насекаде во светот претставува вистински предизвик да се најдат аналитичари на податоци, прво, бидејќи е релативно нова професија, која некогаш бара интердисциплинарни познавања, и второ бидејќи јавниот сектор ја губи битката со приватниот во вработување и задржување на ваквите кадри.

Можности за примена на обемните бази на податоци како алатки за спречување на корупцијата кај нас

За разлика од некои други земји во светот, може да се констатира дека кај нас оваа идеја е во самиот зародиш, иако некои почетни чекори, повеќе околу воспоставување предуслови, веќе се прават.

Единствената дефиниција на овој поим засега ја дава Дирекцијата за заштита на личните податоци во својот Водич за големи податоци и обработка на податоци. „Големите податоци ги претставуваат информативните средства кои се карактеризираат со таква голема количина, брзина и разновидност за да бараат специфични технолошки и аналитички методи за нејзино трансформирање во вредност“.

Сите други спомнувања на поимот кај нас се или во образовен или во бизнис контекст, односно како предмети во наставни програми или како понуди на фирми до своите клиенти.

Поголемо познавање и препознавање на улогата во спречување на корупцијата засега има за отворените податоци. Во оваа смисла, во владината Стратегија за отворени податоци (2018-2020) уште во воведот се препознава дека „отворените податоци можат да бидат силен поттик... за подобрување на управувањето преку зголемување на транспарентноста, отвореноста, интегритетот и јавното учество во централната и локалната власт.“ Натаму, се истакнува дека остварувањето на целта на Стратегијата, меѓу другото, ќе доведе и до „поддржување на напорите за борба против корупцијата...“.

Како и да е, истите можности (заедно со предизвиците) за примена на обемните бази на податоци ѝ стојат на располагање и на Северна Македонија. Во земјата има мноштво бази на податоци кои можат да се стават во функција на спречување на корупцијата и нејзино откривање. Уште повеќе, ако овие обемни бази се поврзат меѓу себе и ако се овозможи вкрстено користење на податоците и нивно споредување.

Во оваа смисла станува збор за податоците на Управата за јавни приходи, на Електронскиот систем за јавни набавки, на Централниот регистар, на Централниот депозитар за хартии од вредност, на Трезорот, на Катастарот на недвижности, на Фондот за пензиско и инвалидско осигурување, на Фондот за здравствено осигурување, на Министерството за информатичко општество и администрација, на Царинската управа, на Министерството за внатрешни работи, на Државната комисија за спречување на корупцијата, Државниот завод за статистика, Државната изборна комисија, итн.

Сите овие институции поседуваат обемни бази на податоци со чија поголема достапност и вкрстување ќе може полесно да се проверуваат, на пример, објективноста на податоците од анкетните листови за имотната состојба на избраните и именуваните функционери; злоупотреби во добивањето на државните тендери; скриени сопственици на фирмите; евентуални неправилности при добивање државна и социјална помош, субвенции и сл.

Чекор напред во оваа смисла беше направен во новиот Закон за спречување на корупцијата и судирот на интереси со воведување на одредба (член 25) во која на Државната комисија за спречување на корупцијата ѝ се овозможува директен електронски пристап и бесплатно користење податоци од базите на податоци кои се водат во дури 17 други органи и институции. (Намерно или случајно, меѓу сиве овие бази на податоци не се вклучени и податоците од Електронскиот систем за јавни набавки.)

Иако, ни самата Државна комисија за спречување на корупцијата засега не ја користи оваа законска поволност и моќна алатка за борба против корупцијата (со образложение дека

нема соодветен простор за сместување на потребната компјутерска опрема која ќе го овозможи користењето на сите овие бази на податоци), сепак, изостанува и пошироката достапност на овие бази (барем на основните) и за другите општествени чинители во борбата против корупцијата во земјава, пред сè граѓаните, граѓанските организации и медиумите.

Недостапноста на обемните бази на податоци ги оневозможуваат и другите институции вклучени во борбата против корупцијата да ја вршат ефикасно својата работа и законските обврски.

Исто така, најголем дел и од оние малкуте достапни бази на податоци не се во отворен формат, со што е оневозможено нивно лесно користење и споредување. Во многу случаи, како оправдување се користи заштитата на лични податоци, иако за ова прашање постои законска регулатива и можности за достапност на податоците со истовремена заштита на оние податоци кои не смеат јавно да се објавуваат. Натаму, како често оправдување за недостапност на базите на податоци е немањето соодветен технички и кадровски потенцијал кај институциите. И ова оправдување може да се надмине со ангажирање на надворешни лица и со набавка на соодветната опрема.

Иако се прават одредени чекори (повеќе на ад-хок ниво) кон достапност на нови бази на податоци (објавување податоци за извршување на Буџетот; заложби во Националниот акциски план за отворено владино партнерство 2018-2020; проценка на ризици во јавните набавки и сл.), сепак може со право да се констатира дека отворањето на базите на податоци, нивната достапност и креирањето на обемни бази на податоци, а со сето тоа и нивното користење во антикорупциски цели оди бавно, неорганизирано и без особена политичка воља за забрзување.

Заклучоци и препораки

Имајќи го сето погоре кажано предвид, може да се заклучи дека целосната примена на анализата на обемните бази на податоци за антикорупциски цели допрва претстои.

Насекаде во светот, од Европа, преку Индија, САД и Кореја, има бројни успешни примери за спречување на корупцијата преку овозможување на обемните дигитални бази да го „прочитаат“ тоа што не може да го стори човекот со обичната опрема со којашто работи секојдневно.

Како што овие бази со огромна брзина и обем ќе се полнат, така ќе растат и нивните антикорупциски можности. Истовремено ќе растат и предизвиците – пред сè менаџирање

со базите, соодветна заштита на личните податоци, наоѓање и задржување на стручни лица за работа со базите и анализа на податоците и сл.

Неисцрпни се можностите за примена на анализата на обемните бази на податоци кај нас имајќи предвид дека земјата се наоѓа на самиот почеток од примената на оваа анализа. Некои пионерски чекори се веќе направени, но ќе биде потребно значително забрзување бидејќи земјата нема да остане имуна ниту на дигиталната револуција, ниту на корупцијата.

За почеток, потребно е да профункционира користењето на сите бази на податоци што треба да ѝ се достапни на Државната комисија за спречување на корупцијата, како прв позначаен пример за примена на обемните бази на податоци за спречување на корупцијата. Ова ќе биде важно особено за спроведување на новата Национална стратегија за спречување на корупцијата и судирот на интереси (2020-2024).

Покрај ова, потребно е целосно да се спроведе Стратегијата за отворени податоци и институциите да ги објавуваат своите бази на податоци во отворен формат. Ова ќе овозможи и поголема транспарентност, отчетност, одговорност и доверба во институциите, и поголема вклученост на граѓаните, и поефикасни активности и кај другите актери во антикорупциската борба – граѓанските организации, медиумите, па и бизнис-секторот.

Во оваа насока треба да се забрза и омасови процесот на дигитализација на услугите, ставањето на располагање на сите масовни бази на податоци и нивно поврзување. Јавниот сектор треба да најде начини за кадровско екипирање со соодветни стручни лица за работа со обемни бази на податоци и за анализа на тие податоци. Секако, овде спаѓа и соработка и интеракција во оваа смисла и со бизнис-секторот.

Профил на авторот

Герман Филков е аналитичар и истражувач во областите добро владеење, транспарентност и јавно трошење, како и ментор на новинари кои истражуваат и известуваат за корупцијата и за јавните трошења. Претседател е на невладината организација Центар за граѓански комуникации од Скопје. По 12 години минати како телевизиски новинар и уредник известувајќи за најзначајните промени во македонското општество, во последните 15 години го применува своето искуство во работењето во граѓанското општество. Има спроведено бројни истражувања и преку изработените анализи и документи за јавни политики има предложено стотици препораки за мерки за стеснување на просторот за корупција и за зголемување на транспарентноста, отчетноста и интегритетот. Редовно учествува во подготовка на антикорупциски стратегии и

иницијативи. Посетувал бројни меѓународни обуки за добро владеење, антикорупција, креирање документи за јавни политики и новинарство. Има дипломирано, магистрирано и докторирано на економски науки, а како визитинг истражувач има специјализирано комуникации на Универзитетот на Мисури во Колумбија, САД.

Библиографија/користена литература

“Analytics for Integrity: Data-Driven Approaches for Enhancing Corruption and Fraud Risk Assessments, OECD, 2019

„Водич за големи податоци и обработка на податоци“, Дирекција за заштита на личните податоци, 2018

“G20 Anti-Corruption Working Group Action Plan 2019-2021”, OECD, 2018

„Закон за спречување на корупцијата и судирот на интереси“, Службен весник на РМ, бр. 12 од 19 јануари 2019 година

“Logos, mythos and ethos of blockchain: An Integrated Framework for Anti-corruption”, Ziya Aliyev, Erasmus University Rotterdam, Iqbal Safarov, School of Governance, Utrecht University, March 2019

“More data, less corruption?”, Global Integrity, Alan Hudson, December 21, 2018

“Research on Anti-Corruption Theory under Big Data Environment”, International Journal of Educational Planning and Administration, 2018

„Стратегија за отворени податоци 2018-2020“, Влада на Република Македонија, Министерство за Информатичко општество и администрација, 2018

“The measurement and macro-relevance of corruption: A big data approach, IMF Working Paper, Sandile Hlatshwayo, Anne Oeking, Manuk Ghazanchyan, David Corvino, Ananya Shukla and Lamin Leigh, August 2018

“Using Data and Transparency to Fight Corruption in Public Procurement”, World Bank Group

“How data is changing the fight against corruption”, Carlos Santiso, OECD-Forum, 2018

“Can we use big data analytics as an anti-corruption tool?”, Michael Dixon, March 15, 2019

“Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Towards a common European data space”, European Commission, 2018

“Compendium of good practices on the use of open data for Anti-corruption: Towards data-driven public sector integrity and civic auditing”, OECD

“Connecting the Dots: Building the Case for Open Data to Fight Corruption”, WVV and Transparency International, 2017