

Številka: 018-077/2026-7

Datum: 28. 7. 2026

SKLEP

Državna revizijska komisija za revizijo postopkov oddaje javnih naročil (v nadaljevanju: Državna revizijska komisija) je na podlagi 39. in 70. člena Zakona o pravnem varstvu v postopkih javnega naročanja (Uradni list RS, št. 43/2011 s sprem.; v nadaljevanju: ZPVPJN) v senatu Aleksandra Petrovčiča kot predsednika senata ter Marka Medveda in Igorja Luzarja kot članov senata v postopku pravnega varstva pri oddaji javnega naročila »*Dostop in konfiguracija dostopov do platforme z vrhunskimi jezikovnimi modeli*«, na podlagi zahtevka za revizijo, ki ga je vložila družba TELEKOM SLOVENIJE, d.d., Cigaletova ulica 15, Ljubljana (v nadaljevanju: vlagatelj), zoper ravnanje naročnika Ministrstvo za izobraževanje, znanost in mladino, Masarykova cesta 16, Ljubljana in pooblaščenega naročnika Ministrstvo za notranje zadeve in javno upravo, Štefanova ulica 2, 1000 Ljubljana (v nadaljevanju: naročnik), 28. 7. 2026

odločila:

1. Zahtevku za revizijo se delno ugotovi tako, da se v dokumentaciji v zvezi z oddajo javnega naročila za tretjo fazo v dokumentu »*Tehnične specifikacije*« razveljavi zahtevo, določeno v četrti alineji točke 5.1.2, ki se glasi:
»[M] jezikovni model (za zahtevano stopnjo 1 - najzmogljivejši) dosega vsaj ≥ 1440 točk na Leaderboard Overview <https://arena.ai/leaderboard-> (Text Arena - Overall (style-controlled))«.

V preostalem delu se zahtevke za revizijo zavrne kot neutemeljen.

2. Naročnik je dolžan vlagatelju povrniti stroške pravnega varstva v višini 2.000,00 EUR v roku 15 dni od prejema tega sklepa, po izteku tega roka pa z zakonskimi zamudnimi obrestmi do plačila. Višja stroškovna zahteva vlagatelja se zavrne.
3. Zahteva naročnika za povrnitev stroškov predrevizijskega postopka se zavrne.

Obrazložitev:

Obvestilo o predmetnem javnem naročilu, ki ga naročnik oddaja v postopku s konkurenčnim dialogom po 42. členu Zakona o javnem naročanju (Uradni list RS, št. 91/2015 s sprem.; v nadaljevanju: ZJN-3), je bilo 11. 11. 2025 objavljeno na portalu javnih naročil pod št. objave JN008809/2025-EUe16/01 in v Dodatku k Uradnem listu Evropske unije pod št. objave 747121-2025.

Naročnik je 6. 2. 2026 na portalu javnih naročil objavil »*Odločitev o priznanju sposobnosti*« št. 430-190/2025-3360-16 z dne 2. 2. 2026 (v nadaljevanju: odločitev o priznanju sposobnosti), s katero je odločil, da se šestim od sedmih kandidatov, ki si oddali prijave (med drugim tudi vlagatelju), prizna sposobnost ter da se te kandidate v nadaljevanju povabi k udeležbi v konkurenčnem dialogu.

Vlagatelj je 28. 5. 2026 vložil zahtevek za revizijo zoper določbe dokumentacije v zvezi z oddajo javnega naročila v tretji fazi postopka, v katerem zatrjuje nezakonitost zahteve, s katero je naročnik določil, da mora jezikovni model (za zahtevano stopnjo 1 – najzmogljivejši) dosegati vsaj 1440 točk na Leaderboard Overview (Text Arena – Overall (style-controlled), t. i. lestvici Arena. Zahteva tudi povrnitev priglašениh stroškov postopka pravnega varstva.

Vlagatelj najprej zatrjuje, da je naročnik tehnično zahtevo, po kateri mora model stopnje 1 dosegati najmanj 1440 točk na lestvici Arena, oblikoval na podlagi neustreznih in subjektivnih kriterijev. Pojasnjuje, da Arena temelji na anonimnem glasovanju uporabnikov, ki med odgovori različnih modelov izbirajo tistega, ki se jim zdi boljši, zato rezultati ne odražajo objektivne tehnične zmogljivosti modelov, temveč agregat subjektivnih preferenc uporabnikov. Po njegovem mnenju uporabniki ne ocenjujejo pravilnosti, zanesljivosti ali varnosti odgovorov, ampak predvsem njihovo všečnost, slog, prepričljivost in uporabniško izkušnjo. Ob tem poudarja, da lestvica Arena ne temelji na standardiziranih testih, ponovljivih rezultatih ali enotnih vhodnih podatkih, zato naj ne bi omogočala objektivne primerjave modelov. Ker se pri izračunu upoštevajo glasovi uporabnikov skozi daljše časovno obdobje, naj rezultat posameznega modela ne bi odražal njegove trenutne kakovosti, temveč preteklo uspešnost. Posledično se izboljšave modelov ne odražajo nujno v njihovem trenutnem številu točk, medtem ko lahko nekateri modeli zaradi preteklih pozitivnih ocen ohranjajo visoko uvrstitev kljub morebitnemu poslabšanju kakovosti. Po mnenju vlagatelja Arena meri predvsem kakovost pogovora, slog komunikacije in splošni vtis uporabnika, ne pa lastnosti, ki so bistvene za predmet naročila (kot so varnost, evropska podatkovna suverenost, produkcijska razširljivost, API-integracije, podpora agentnim funkcionalnostim, upravljanje uporabnikov, zanesljivost infrastrukture in kakovost delovanja v slovenskem jeziku), zato naj ne bi predstavljala ustreznega dokazila o izpolnjevanju tehničnih zahtev naročila. Vlagatelj posebej izpostavlja model Mistral Large 3, ki po njegovih navedbah s 1414 točkami dosega nižjo uvrstitev (90. mesto) na lestvici Arena iz razloga, ker daje prednost tehničnim, jedrnatim in manj pogovorno usmerjenim odgovorom, medtem ko Arena favorizira modele z daljšimi, bolj prijaznimi in uporabnikom bolj všečnimi odgovori. Po njegovem mnenju je Mistral Large 3 na številnih drugih strokovnih lestvicah ocenjen bistveno bolje, zlasti glede tehnične zmogljivosti, sklepanja, programiranja, poslovne uporabe in razmerja med zmogljivostjo ter ceno. Kot primer vlagatelj navaja lestvico n8n AI Benchmark, ki po njegovem mnenju bolje odraža dejanske produkcijske zahteve naročila, saj ocenjuje zanesljivost modelov pri avtomatizacijah, delu z orodji, strukturiranih odgovorih in delovanju agentov umetne inteligence. Meni, da Arena meri le ozek segment sposobnosti modelov in zato ni primerna kot obvezna tehnična zahteva v predmetnem postopku javnega naročanja.

Vlagatelj nadalje zatrjuje, da naročnik ni pojasnil, zakaj je prag določil ravno pri 1440 točkah, saj iz dokumentacije ni razvidna metodologija določitve tega praga, prav tako ni pojasnjeno, zakaj model s 1414 točkami ni tehnično ustrezen, model s 1440 točkami pa je. Zahteva naj bi bila zato po mnenju vlagatelja arbitrarna, netransparentna in brez objektivne povezave s funkcionalnimi potrebami naročnika. Navaja, da je naročnik v postopku konkurenčnega dialoga poznal modele, ki jih ponujajo posamezni kandidati, nato pa je po izvedenem dialogu določil prag 1440 točk, s čimer naj bi neupravičeno omejil konkurenco in izločil model Mistral Large 3, ki ga vlagatelj namerava ponuditi. Po mnenju vlagatelja gre tako za nesorazmerno zahtevo, ki ni povezana z dejanskimi tehničnimi, varnostnimi in infrastrukturnimi zahtevami predmeta naročila, zato

zaključuje, da zahteva po najmanj 1440 točkah na lestvici Arena ne predstavlja dopustne tehnične specifikacije v smislu 68. člena ZJN-3, ter da ni objektivno utemeljena niti sorazmerna s predmetom naročila, prav tako pa neupravičeno omejuje konkurenco in vlagatelju onemogoča oddajo dopustne ponudbe.

Naročnik je na podlagi tretjega odstavka 17. člena ZPVPJN s sklepom z dne 28. 5. 2026 zadržal izvajanje predmetnega postopka oddaje javnega naročila do pravnomočne odločitve naročnika ali Državne revizijske komisije o zahtevku za revizijo.

Naročnik je s sklepom št. 430-190/2025-3360-28 z dne 9. 6. 2026 (v nadaljevanju: odločitev o zahtevku za revizijo) zavrnil vlagateljev zahtevek za revizijo in njegovo zahtevo za povrnitev stroškov postopka ter mu naložil povračilo stroškov, ki so naročniku nastali v predrevizijskem postopku.

Uvodoma naročnik v okviru predhodnega preizkusa zahtevka za revizijo opozarja na vprašanje obstoja oziroma uporabe omejitve iz tretjega odstavka 16. člena ZPVPJN. Navaja, da se vlagatelj v zahtevku sklicuje na vprašanje, ki ga je prek sistema e-JN postavil 21. 5. 2026 in v katerem je vprašal, ali naročnik poleg predvidenega ocenjevanja kakovosti modelov po lestvici Arena dopušča uporabo dodatnih oziroma alternativnih mehanizmov rangiranja oziroma ocenjevanja, za katere vlagatelj meni, da bolje odražajo dejansko zmogljivost modelov (naročnik pa je na to vprašanje odgovoril, da je predvidel uporabo lestvic Arena in da alternativnih mehanizmov ne dopušča). Po naročnikovem stališču iz navedenega vprašanja ne izhaja opozorilo o nezakonnosti (neustreznosti) lestvice Arena niti opozorilo o spornosti zahtevanega praga 1440 točk, temveč se nanaša zgolj na možnost uporabe dodatnih oziroma alternativnih lestvic. Naročnik zato meni, da vlagatelj ni jasno in konkretno opozoril na kršitve, ki jih sedaj uveljavlja v zahtevku za revizijo, posledično pa je v postopku pravnega varstva mogoče presojati kvečjemu vprašanje, ali bi moral naročnik dopustiti alternativne mehanizme rangiranja ponujenih modelov, ne pa tudi vprašanja ustreznosti Arena lestvice kot take ali zakonitosti praga 1440 točk. Ker vlagatelj v fazi vprašanj teh očitkov ni izpostavil, naročnik meni, da zanje niso izpolnjene procesne predpostavke za vsebinsko obravnavo po 16. členu ZPVPJN, kljub temu pa se v nadaljevanju iz previdnosti vsebinsko opredeljuje do vseh vlagateljevih očitkov.

Naročnik pojasnjuje, da je zahtevo glede doseganja določenega praga na lestvici Arena določil premišljeno in v povezavi s predmetom naročila, katerega namen je zagotoviti dostop do jezikovnih modelov za približno 270.000 uporabnikov s področja izobraževanja, raziskovanja in razvoja. Ker je uporaba platforme prostovoljna, je za uspeh naročila ključno, da so ponujeni modeli za uporabnike privlačni, uporabni in prilagojeni njihovim različnim potrebam. Zato je kot pomemben pokazatelj kakovosti izbral prav lestvico Arena, ki temelji na preferencah dejanskih uporabnikov in po njegovem mnenju najbolje odraža splošno uporabnost jezikovnih modelov. Ob tem poudarja, da je pri določanju tehničnih zahtev avtonomen ter da sam presoja, katere lastnosti predmeta naročila zanj predstavljajo ekonomsko korist. Po njegovem mnenju vlagatelj neupravičeno poskuša nadomestiti naročnikovo presojo s svojo oceno, katere lastnosti bi morale biti pomembnejše. Poudarja, da za njegove potrebe ni odločilna zgolj matematična natančnost, kakovost slovenskega jezika, zmogljivost API-jev ali druge posamezne tehnične lastnosti, temveč predvsem splošna uporabnost modelov za širok krog uporabnikov. Ker ni enotnih objektivnih kriterijev za določitev »najboljšega« jezikovnega modela, je izbira meril odvisna od potreb naročnika. Po naročnikovem stališču lestvica Arena predstavlja primeren in objektivni instrument za ocenjevanje kakovosti modelov, saj temelji na milijonih anonimnih primerjav odgovorov različnih modelov. Posamezni uporabniki ne vedo, kateri model je pripravil posamezen odgovor, subjektivnost posameznih ocen pa se zaradi velikega števila glasov statistično izravna. Lestvica temelji na javno objavljeni metodologiji in se za vse ponudnike uporablja enako, pri čemer bo naročnik upošteval vnaprej določen posnetek lestvice z dne 6. 5. 2026, ki je sestavni del

dokumentacije. Naročnik dalje zavrača očitke, da lestvica Arena nagraduje predvsem slogovno bogatejše ali daljše odgovore. Izpostavlja, da je uporabil različico lestvice s t. i. »style-controlled« metodologijo, ki odpravlja vpliv dolžine, oblike in sloga odgovorov ter ocenjuje dejanske sposobnosti modelov. Prav tako meni, da očitki glede spreminjanja rezultatov skozi čas niso relevantni, saj se zahteva preverja glede na stanje lestvice na vnaprej določen presečni datum. Morebitna časovna nihanja po njegovem mnenju ne vplivajo na primerljivost, enakopravnost ali preverljivost ocenjevanja. Naročnik nadalje poudarja, da Arena ni edina zahteva v razpisni dokumentaciji, temveč le eden od elementov presoje kakovosti jezikovnih modelov. Druge lastnosti, na katere opozarja vlagatelj (npr. varnost, obdelava podatkov v EGP, podatkovna rezidenca, integracije API, razširljivost, upravljanje uporabnikov in raven storitve), so urejene z drugimi tehničnimi zahtevami. Enako velja za kakovost delovanja v slovenskem jeziku, za katero je naročnik določil ločene minimalne zahteve in samostojno metodologijo preverjanja, zato se ta ne ocenjuje prek lestvice Arena. Po mnenju naročnika je tako predmetna zahteva povezana s predmetom naročila, objektivno opravičljiva, strokovno utemeljena in sorazmerna njegovim potrebam. Vlagatelj pa po njegovem stališču ni izkazal, da bi bila zahteva diskriminatorna ali da bi mu onemogočala sodelovanje v postopku, saj ni dokazal, da ne razpolaga z drugim modelom, ki bi zahtevo lahko izpolnil. Naročnik zato meni, da vlagateljeve navedbe temeljijo predvsem na nestrinjanju z naročnikovo presojo potreb in ne izkazujejo nezakonitosti sporne tehnične zahteve. Naročnik še pojasnjuje, da ni dopustil uporabe alternativnih lestvic, ker bi to onemogočilo primerljivost ponudb in preverljivost izpolnjevanja zahtev. Uporaba ene same javne, neodvisne in vnaprej določene lestvice po njegovem mnenju zagotavlja enakopravno obravnavo vseh ponudnikov ter omogoča objektivno primerjavo ponujenih modelov.

Glede očitka v zvezi z določitvijo praga 1440 točk naročnik uvodoma pojasnjuje, da ni dolžan že v razpisni dokumentaciji posebej utemeljevati razlogov za določitev posameznih minimalnih tehničnih zahtev, vendar so bile te kljub temu določene na podlagi njegovih potreb. Pojasnjuje, da je za tri zahtevnostne stopnje jezikovnih modelov določil minimalne pragove 1440, 1380 in 1300 točk na lestvici Arena, pri čemer je prag 1440 točk vezan na model stopnje 1 (najzmogljivejši), namenjen najzahtevnejšim primerom uporabe. Dalje navaja, da je pragove oblikoval v sodelovanju z akademsko skupnostjo in upoštevajoč uveljavljene teorije minimalnih standardov. Na tej podlagi je modele razvrstil v štiri kakovostne razrede: »odlične« (približno najboljših 15 % modelov), »dobre« (do približno 40 %), »zadostne« (do približno 66 %) in »nezadostne«. Ker je bilo na lestvici Arena razvrščenih 356 modelov, je prag 1440 točk ustrezal približno prvim 51 modelom (oziroma najboljšim 15 %), prag 1380 točk prvim 135 modelom, prag 1300 točk pa prvim 228 modelom. Po mnenju naročnika je zahteva po najmanj 1440 točkah neposredno povezana s predmetom naročila, saj zagotavlja, da bo najzahtevnejši kakovostni razred zapolnjen z modeli iz dejanskega vrha zmogljivosti. Naročnik poudarja, da razlika med modeli nad in pod pragom ni zgolj formalna ali kozmetična, temveč vsebinska. Nižji prag bi po mnenju naročnika omogočil, da bi ponudniki za najzmogljivejši razred ponudili cenejše in manj zmogljive modele, čeprav razpolagajo z bistveno boljšimi. Glede vlagateljevega modela Mistral Large 3 naročnik izpostavlja, da model s 1414 točkami ni le 26 točk pod pragom, temveč je na lestvici Arena uvrščen šele na 90. mesto, med njim in pragom pa je še 38 modelov. Zato po njegovem mnenju ne gre za zanemarljivo razliko, temveč za statistično in vsebinsko razliko v zmogljivosti. Model vlagatelja se po naročnikovi metodologiji uvršča v kakovostni razred »dobro«, medtem ko je za stopnjo 1 zahtevan razred »odlično«. Naročnik tako zavrača očitke, da prag 1440 točk favorizira ozko skupino modelov ali posamezne ponudnike. Navaja, da je na presečni datum prag dosegalo 51 modelov dvanajstih različnih ponudnikov, zato zahteve ni mogoče šteti za diskriminatorno ali prilagojeno posameznemu ponudniku. Ker je za stopnjo 1 zahtevan le en tak model, naročnik meni, da zahteva zagotavlja zadostno konkurenco. Po naročnikovem stališču je določitev minimalne ravni kakovosti v njegovi pristojnosti, če izkaže, da modeli pod pragom ne ustrezajo njegovim potrebam in da zahteva ne omejuje konkurence neupravičeno – ker je konkretnem primeru nabor več kot 50 ustreznih modelov dovolj širok, po mnenju naročnika ni

mogoče govoriti o neupravičenem omejevanju konkurence, temveč o objektivni meji tehnične primernosti. Naročnik tudi na tem mestu zavrača vlagateljevo sklicevanje na lestvico n8n. Pojasnjuje, da Arena meri splošno uporabniško izkušnjo in kakovost odgovorov v realnih pogovorih ter je zato primerna za potrebe njegovih 270.000 uporabnikov, medtem ko je n8n specializirana lestvica za avtomatizacijske delovne tokove, ki za predmet naročila ni bistvena. Poleg tega naj bi bila lestvica n8n bistveno ožja, saj zajema le okoli 60 modelov. Naročnik še poudarja, da se tudi po tej lestvici modeli Mistral ne uvrščajo v najvišji kakovostni razred, zato vlagatelj ni izkazal, da bi bili na drugih lestvicah bistveno bolje ocenjeni. Naročnik nadalje navaja, da je bil pomen lestvice Arena ponudnikom predstavljen že v fazi dialoga, hkrati pa so lahko ponudniki v tretji fazi ponudili tudi druge modele od tistih, ki so jih navedli v prijavi. Ker so bili izpisi iz Arena lestvice vključeni v dokumentacijo, so lahko ponudniki vnaprej preverili, kateri modeli izpolnjujejo zahteve, in temu prilagodili svoje ponudbe.

Naročnik sklepno poudarja, da je prag 1440 točk povezan s predmetom naročila, strokovno utemeljen in sorazmeren, ker ga dosegajo veliko število modelov različnih ponudnikov, pa tudi ne omejuje konkurence. Izločitev vlagateljevega modela je tako po mnenju naročnika posledica njegove dejanske uvrstitve na lestvici in ne arbitrarno postavljene zahteve, zato ocenjuje, da sporna tehnična zahteva izpolnjuje pogoje iz 68. člena ZJN-3 ter zagotavlja enakopravno obravnavo ponudnikov.

Vlagatelj se je do navedb naročnika v odločitvi o zahtevku za revizijo opredelil z vlogo z dne 12. 6. 2026, v kateri prereka naročnikove razloge za zavrnitev zahtevka za revizijo.

Vlagatelj najprej nasprotuje stališču naročnika, da je v tretji fazi postopka opozorilo prek sistema e-JN procesni pogoj za vsebinsko obravnavo zahtevka za revizijo. Poudarja, da ZPVPJN kot procesno predpostavko določa izključno opozorilo prek portala javnih naročil, ne pa opozorila prek sistema e-JN. Ker portal javnih naročil in e-JN opravljata različni funkciji, po njegovem mnenju opozorila prek e-JN ni mogoče enačiti z opozorilom prek portala javnih naročil. Posledično po njegovem stališču odsotnost takšnega opozorila v tej fazi ne more preprečiti vsebinske obravnave zahtevka za revizijo. Ne glede na navedeno nadalje zavrača naročnikovo stališče, da njegovo vprašanje glede na vsebino ni predstavljalo opozorila na zatrevano kršitev. Sklicuje se na ustaljeno prakso Državne revizijske komisije, iz katere izhaja, da zakon ne določa posebne oblike ali minimalne vsebine opozorila ter da je treba pogoj opozorila razlagati široko. Kot opozorilo je zato mogoče šteti tudi vprašanja, pripombe ali predloge, če se po vsebini nanašajo na očitke, ki se kasneje uveljavljajo v zahtevku za revizijo. Glede na navedeno vlagatelj meni, da vprašanje, v katerem je navedel, da alternativni mehanizmi rangiranja oziroma ocenjevanja bolje odražajo dejansko zmogljivost jezikovnih modelov kot lestvica Arena, vsebinsko predstavljalo opozorilo na neustreznost uporabe lestvice Arena kot merila za ocenjevanje zmogljivosti modelov. Ker se zahteve glede doseganja določenega števila točk na lestvici Arena, vključno z zahtevo po najmanj 1440 točkah, nahajajo v istem sklopu tehničnih specifikacij, je po njegovem stališču iz vprašanja jasno razvidno, da dvomi v ustreznost izbranega mehanizma ocenjevanja in posledično tudi v zakonitost zahtev, ki na njem temeljijo. Vlagatelj zato meni, da je naročnika jasno, konkretno in nedvoumno opozoril na očitano kršitev, čeprav tega po njegovem stališču sploh ni bil dolžan storiti, saj v tej fazi postopka možnost opozorila prek portala javnih naročil ni obstajala. Ob tem še izpostavlja, da je naročnik zahtevek za revizijo vsebinsko obravnaval in se opredelil do vseh zatrevanih kršitev. Po njegovem mnenju bi moral naročnik, če bi štel, da procesna predpostavka iz tretjega odstavka 16. člena ZPVPJN ni izpolnjena, zahtevek za revizijo zavreči, česar pa ni storil.

Vlagatelj dalje vztraja, da naročnik zahteve po vrhunskem oziroma najzmogljivejšem jezikovnem modelu ne more utemeljiti z rezultatom na lestvici Arena, saj ta ne meri objektivnih tehničnih ali funkcionalnih lastnosti modelov, temveč zgolj preference uporabnikov. Po njegovem mnenju

vrhunska kakovost modela ne more biti določena na podlagi priljubljenosti oziroma všečnosti odgovorov, temveč na podlagi preverljivih tehničnih lastnosti in funkcionalnosti, ki so neposredno povezane s predmetom naročila. V zvezi s tem navaja primerjavo z avtomobilom, pri katerem bi bila ustrezna tehnična zahteva varnostna ocena, ne pa uvrstitev na lestvici priljubljenosti. Poudarja, da naročnik ni izkazal, zakaj naj bi rezultat na lestvici Arena dokazoval primernost modela za potrebe slovenskega izobraževalnega in raziskovalnega okolja, saj ni pojasnil, ali so uporabniki, katerih preference oblikujejo lestvico Arena, primerljivi z njegovimi ciljnim uporabniki (učenci, študenti, učitelji in raziskovalci), niti ni dokazal, da njihove preference odražajo potrebe teh skupin. Po mnenju vlagatelja naročnik zmotno enači uporabniške preference z objektivnimi tehničnimi lastnostmi modelov. Dejstvo, da uporabniki ne vedo, kateri model ocenjujejo, in da pri ocenjevanju sodeluje veliko število uporabnikov, po njegovem stališču ne spremeni dejstva, da Arena ostaja sistem preferenčnega glasovanja, ki meri subjektivno zaznano kakovost odgovorov in ne njihove pravilnosti, zanesljivosti ali tehnične ustreznosti. Tudi različica »style-controlled« naj bi zgolj zmanjšala vpliv sloga odgovorov, ne pa omogočila objektivnega merjenja tehnične kakovosti modelov. Vlagatelj nadalje izpostavlja, da naročnik ni izkazal neposredne povezave med rezultatom na lestvici Arena in dejanskimi funkcionalnimi zahtevami predmeta naročila, saj ni pojasnil, kako višje število točk na lestvici vpliva na učinkovitost modela v izobraževanju in raziskovanju, niti kako naj bi se to odražalo v boljšem delovanju, večji kakovosti storitve ali ekonomskih koristih. Nasprotno naj bi naročnik predmet naročila neutemeljeno zreduciral na splošno uporabniško zadovoljstvo z odgovori modelov. Vlagatelj opozarja tudi, da uporaba enotnega presečnega datuma in enakega posnetka lestvice za vse ponudnike zagotavlja le formalno enakopravno obravnavo, ne dokazuje pa, da lestvica Arena predstavlja ustrezno merilo tehnične kakovosti – model, ki je tehnično izboljššan, lahko zaradi zgodovinskih rezultatov še vedno ne dosega zahtevanega praga, medtem ko lahko drug model zahtevo izpolnjuje na podlagi preteklih ocen. Dejstvo, da je naročnik določil še ločeno preverjanje sposobnosti delovanja v slovenskem jeziku, po stališču vlagatelja dodatno potrjuje, da Arena ne predstavlja celovitega ali zadostnega pokazatelja kakovosti modelov. Zato bi moral naročnik posebej dokazati, zakaj je ravno ta lestvica primerna podlaga za določitev minimalnih tehničnih zahtev. Ob tem poudarja, da ne nasprotuje naročnikovi pravici določati tehnične zahteve, temveč zatrjuje, da naročnik ni izkazal njihove objektivne povezanosti s predmetom naročila, sorazmernosti in nujnosti. Prav tako meni, da naročnik neupravičeno zavrača možnost uporabe enakovrednih dokazil, čeprav bi moral ob sklicevanju na določeno ocenjevalno lestvico omogočiti tudi dokazovanje enakovrednosti z drugimi metodologijami.

Glede praga 1440 točk vlagatelj navaja, da naročnik ni dokazal njegove strokovne utemeljenosti, temveč zgolj pojasnil, da ustreza vnaprej izbranemu razredu najboljših 15 % modelov na lestvici Arena. Po njegovem mnenju zato prag ne izhaja iz dejanskih potreb naročnika, temveč iz arbitrarno določene statistične razvrstitve modelov. Če bi bilo na lestvici več ali manj modelov oziroma drugačna porazdelitev rezultatov, bi bil tudi prag drugačen, čeprav bi potrebe naročnika ostale enake. Vlagatelj izpodbija tudi naročnikovo metodologijo ocenjevanja modelov glede na njihovo razvrstitev na lestvici Arena (odlično, dobro, zadostno). Meni, da tak pristop ne dokazuje, da modeli pod pragom 1440 točk niso sposobni izpolnjevati najzahtevnejših nalog, niti da modeli nad pragom to nujno zmorejo bolje. Naročnik naj bi zato relativno uvrstitev modelov pretvoril v absolutni tehnični pogoj brez ustrezne strokovne podlage. Po mnenju vlagatelja naročnik neutemeljeno enači višjo uvrstitev na lestvici Arena z večjo tehnično sposobnostjo modela. Dejstvo, da je model Mistral Large 3 uvrščen na 90. mesto namesto med prvih 51 modelov, po njegovem stališču ne dokazuje, da ni sposoben izvajati nalog, predvidenih za model stopnje 1. Naročnik naj bi izkazoval le obstoj zadostnega števila modelov nad pragom, ne pa tudi nujnosti in sorazmernosti samega praga. Vlagatelj nadalje poudarja, da različne priznane lestvice modele ocenjujejo različno, kar po njegovem mnenju potrjuje, da Arena ni univerzalno in objektivno merilo kakovosti. Zato naročnik ne more utemeljiti izključujoče zahteve zgolj s sklicevanjem na eno lestvico. Enako zavrača naročnikove navedbe, da bi lahko ponudil drug model, saj vprašanje po

njegovem mnenju ni, ali na trgu obstajajo modeli, ki zahtevo izpolnjujejo, temveč ali je sama zahteva zakonita, objektivno utemeljena in sorazmerna. Sklepno vlagatelj vztraja, da naročnik ni dokazal, da prag 1440 točk predstavlja minimalno raven zmogljivosti, nujno za izvedbo najzahtevnejših primerov uporabe. Ker zahteva temelji na relativni razvrstitvi modelov na lestvici Arena in ne na dokazani minimalni tehnični sposobnosti modelov, po njegovem mnenju ostajata njena objektivna utemeljenost in sorazmernost neizkazani.

Naročnik je Državni revizijski komisiji 12. 6. 2026 in na njen poziv še 6. 7. 2026 odstopil dokumentacijo v zvezi s predmetnim javnim naročilom in dokumentacijo predrevizijskega postopka.

Državna revizijska komisija je skladno z 31. členom ZPVPJN zahtevek za revizijo pred meritorno obravnavo predhodno preizkusila in po uradni dolžnosti preverila obstoj procesnih predpostavk, in sicer: ali je vložen pravočasno in pri naročniku; ali vsebuje vse obvezne sestavine iz 15. člena ZPVPJN; ali ga je vložila aktivno legitimirana oseba iz 14. člena ZPVPJN; ali obstajajo omejitve iz 16. člena ZPVPJN in ali je dopusten.

Kljub temu, da je naročnik vlagatelj zahtevek za revizijo (v celoti) vsebinsko obravnaval, je v okviru predhodnega preizkusa zahtevka za revizijo izpostavil obstoj omejitve iz 16. člena ZPVPJN ter hkrati predlagal, da se do tega vprašanja opredeli tudi Državna revizijska komisija.

Naročnik meni, da ga vlagatelj v sistemu e-JN (v fazi zastavljanja vprašanj) na zatrjevano kršitev ni jasno in konkretno opozoril, saj iz zastavljenega vprašanja izhaja, da prosi za pojasnilo, ali naročnik dopušča uporabo dodatnih (alternativnih) mehanizmov rangiranja oziroma ocenjevanja (za katere vlagatelj meni, da bolj reprezentativno odražajo dejansko zmogljivost jezikovnih modelov), v zahtevku za revizijo pa izpodbija ustreznost lestvice Arena in zahtevo po doseganju vsaj 1440 točk na tej lestvici. Vsebina vprašanja in vsebina zahtevka se tako po mnenju naročnika ne nanašata na isto vsebino, zato ocenjuje, da je v predmetnem postopku mogoče presoјati zgolj vprašanje, ali bi moral naročnik dopustiti uporabo dodatnih oziroma alternativnih mehanizmov rangiranja oziroma ocenjevanja, ne pa tudi vprašanja ustreznosti konkretno določenega mehanizma ocenjevanja, vključno z zahtevanim številom točk.

Vlagatelj na drugi strani zatrjuje, da procesna omejitev iz tretjega odstavka 16. člena ZPVPJN v obravnavanem primeru ni uporabljiva, saj v tretji fazi postopka opozorilo prek portala javnih naročil objektivno ni bilo mogoče, ne glede na to pa meni, da je naročnika na očitane kršitve ustrezno opozoril tudi prek informacijskega sistema e-JN.

Tretji odstavek 16. člena ZPVPJN določa, da se v predrevizijskem in revizijskem postopku ne presoјajo očitane kršitve, ki se nanašajo na vsebino objave, povabilo k oddaji ponudb ali razpisno dokumentacijo, če bi lahko vlagatelj ali drug morebitni ponudnik prek portala javnih naročil naročnika opozoril na očitano kršitev, pa te možnosti ni uporabil. Šteje se, da bi vlagatelj ali drug morebitni ponudnik prek portala javnih naročil lahko opozoril na očitano kršitev, če je bilo v postopku javnega naročanja na portalu javnih naročil objavljeno obvestilo o naročilu, na podlagi katerega ponudniki oddajo prijave ali ponudbe.

Kot je Državna revizijska komisija uvodoma že izpostavila, naročnik predmetni postopek oddaje javnega naročila izvaja po postopku s konkurenčnim dialogom (po 42. členu ZJN-3). Naročnik je po zaključku prve faze, v kateri je usposobljenim gospodarskim subjektom (kandidatom) priznal sposobnost, z njimi nadaljeval dialog, pri čemer je v tej fazi komunikacija med naročnikom in kandidati potekala (izključno) prek informacijskega sistema e-JN. Upoštevajoč navedeno v tej fazi postopka možnost postavljanja vprašanj prek portala javnih naročil ni bila več zagotovljena.

Državna revizijska komisija ne more slediti naročnikovemu stališču, da je treba tudi opozorilo, podano prek informacijskega sistema e-JN, šteti za opozorilo v smislu tretjega odstavka 16. člena ZPVPJN. Takšna razlaga namreč nima podlage v zakonskem besedilu. Kot pravilno opozarja vlagatelj, že iz jezikovne razlage navedene določbe izhaja, da je procesna omejitev izrecno vezana na možnost opozorila prek portala javnih naročil – in torej ne tudi prek informacijskega sistema, ki se uporablja za elektronsko komunikacijo med naročnikom in ponudniki (kot je npr. sistem e-JN). Državna revizijska komisija se dalje strinja z vlagateljem, da ZJN-3 portal javnih naročil in elektronska komunikacijska sredstva (kamor sodi tudi sistem e-JN) ureja z različnim namenom. Medtem ko je portal javnih naročil namenjen javni objavi obvestil v zvezi z javnimi naročili in seznanitvi vseh zainteresiranih gospodarskih subjektov z relevantnimi informacijami (prim. 28. točko prvega odstavka 2. člena, drugi odstavek 6. člena, 52. člen ter prvi in drugi odstavek 67. člena ZJN-3), so elektronska komunikacijska sredstva namenjena predvsem sporočanju in izmenjavi informacij med naročnikom in sodelujočimi gospodarskimi subjekti, zlasti elektronski oddaji ponudb (prim. 37. člen ZJN-3).

Pri razlagi obravnavane določbe pa je treba upoštevati tudi njeno naravo. Tretji odstavek 16. člena ZPVPJN predstavlja procesno omejitev, saj preprečuje vsebinsko obravnavo posameznih revizijskih očitkov. Ker gre za izjemo od splošnega pravila, da Državna revizijska komisija vsebinsko presoja (vse) zatrjevane kršitve, mora biti njen domet jasno in izrecno določen z zakonom. Glede na to, da ZPVPJN navedeno procesno omejitev izrecno veže na možnost opozorila prek portala javnih naročil, ne pa tudi prek drugih komunikacijskih kanalov, predmetne določbe ni mogoče razlagati tako, da bi se njena uporaba razširila tudi na primere, ko zastavljanje vprašanj prek portala javnih naročil v posamezni fazi postopka ni bilo več na voljo ter je komunikacija med naročnikom in gospodarskimi subjekti potekala izključno prek informacijskega sistema e-JN. Takšna razlaga bi namreč pomenila širitev zakonsko določene procesne omejitve preko njenega jasnega besedila. Če bi zakonodajalec želel procesni pogoj vezati tudi na opozorila, podana prek drugih informacijskih sistemov, bi to lahko izrecno določil, vendar tega ni storil.

Državna revizijska komisija tako nima podlage, da bi z (ekstenzivno) razlago širila zakonsko določeno procesno omejitev tudi na primere, ko je komunikacija med naročnikom in omejenim krogom gospodarskih subjektov (tj. kandidati, katerim je naročnik v predhodni fazi priznal usposobljenost) potekala zgolj prek informacijskega sistema e-JN. Takšno razumevanje potrjuje tudi zakonodajno gradivo ob sprejemu novele ZPVPJN-B (Zakon o spremembah in dopolnitvah Zakona o pravnem varstvu v postopkih javnega naročanja, Uradni list RS, št. 60/17). Zakonodajalec je namreč glede na prejšnjo ureditev, po kateri je opustitev opozorila pomenila izgubo aktivne legitimacije, določil celo milejšo posledico, in sicer je z navedeno spremembo procesnega zakona določil, da se zgolj ne presojuje posamezni očitki, glede katerih opozorilo ni bilo podano. Iz tega gre razbrati, da je namen predmetne omejitve iz tretjega odstavka 16. člena ZPVPJN vzpodbujanje predhodne komunikacije med gospodarskimi subjekti in naročnikom oziroma seznanitev naročnika z domnevno kršitvijo že med postopkom oddaje javnega naročila ter posledično predhodna odprava morebitnih nezakonitosti in preprečitev nepotrebnih revizijskih postopkov, ne pa formalno omejevanje dostopa do pravnega varstva.

Upoštevajoč vse navedeno Državna revizijska komisija zaključuje, da v danem primeru niso podani pogoji za uporabo procesne omejitve iz tretjega odstavka 16. člena ZPVPJN.

Iz tega razloga se Državna revizijska komisija ni opredeljevala do navedb strank glede tega, ali je bilo vprašanje, podano prek informacijskega sistema e-JN, dovolj konkretno oziroma ali je po vsebini predstavljalo opozorilo na (vse) zatrjevane kršitve oziroma, saj to za odločitev o obstoju procesne predpostavke v obravnavani zadevi ni bistveno. Navedena presoja bi bila namreč potrebna le v primeru uporabe omejitve iz tretjega odstavka 16. člena ZPVPJN, ki pa v

obravnavanem primeru, kot že pojasnjeno, po presoji Državne revizijske komisije ne pride v poštev.

Ker je Državna revizijska komisija ugotovila, da v konkretnem primeru pomisleki naročnika glede obstoja omejitve iz tretjega odstavka 16. člena ZPVPJN niso utemeljeni, prav tako pa je ugotovila, da so izpolnjeni tudi vsi ostali pogoji iz prvega odstavka 31. člena ZPVPJN, je zahtevek za revizijo na podlagi drugega odstavka 31. člena ZPVPJN (v celoti) sprejela v obravnavo.

Po pregledu dokumentacije o javnem naročilu in dokumentacije, predložene v postopku pravnega varstva, ter po preučitvi navedb vlagatelja in naročnika je Državna revizijska komisija odločila, kot izhaja iz izreka tega sklepa, iz razlogov, navedenih v nadaljevanju.

Med vlagateljem in naročnikom je spor o tem, ali je naročnik skladno z ZJN-3 v dokumentaciji v zvezi z oddajo javnega naročila določil uporabo lestvice Arena za ugotavljanje zmogljivosti ponujenih jezikovnih modelov ter ali je za jezikovni model stopnje 1 zakonito določil minimalni prag najmanj 1440 točk na navedeni lestvici.

V obravnavanem primeru je naročnik že v prvi fazi postopka, v okviru prvotne razpisne dokumentacije (objavljene na portalu javnih naročil z obvestilom o naročilu št. JN008809/2025-EUe16/01 z dne 11. 11. 2025), podal določene tehnične zahteve (dokument »*Okvirna tehnična specifikacija*«). Končne in dopolnjene tehnične zahteve pa je naročnik določil (še) v tretji fazi postopka (po izvedenem dialogu z usposobljenimi kandidati), in sicer v dokumentu »*Tehnične specifikacije*« (v nadaljevanju: tehnične specifikacije), ki predstavlja del dokumentacije v zvezi z oddajo javnega naročila za tretjo fazo (dokument »*Dokumentacija v zvezi z oddajo javnega naročila po postopku konkurenčnega dialoga z oznako KDLLM-1/2025, št. 430-190/2025-3360-24 z dne 6. 5. 2026*«).

Naročnik je med drugim v tehničnih specifikacijah v poglavju 5.1 podal zahteve za platformo in v točki »5.1.2 *Portfelj LLM¹ in zmogljivost*« za vse zahtevane stopnje jezikovnih modelov (1, 2 in 3) za ugotavljanje njihove zmogljivosti določil uporabo lestvice »*Leaderboard Overview (Text Arena – Overall (style-controlled))*« (v nadaljevanju: lestvica Arena) ter navedel zahtevani točkovni prag na tej lestvici za vsako od teh stopenj.

Naročnik je tako za (najzmogljivejšo) stopnjo 1 v četrti alineji točke 5.1.2 tehničnih specifikacij določil:

»[M] jezikovni model (za zahtevano stopnjo 1 - najzmogljivejši) dosega vsaj ≥ 1440 točk na *Leaderboard Overview* <https://arena.ai/leaderboard-> (*Text Arena - Overall (style-controlled)*)«.

Naročnik je dokumentaciji v zvezi z oddajo javnega naročila za tretjo fazo priložil tudi dokument »*Izpisi in zaslonke slike LM Arena*«, ki prikazuje zaslonke posnetke lestvice Arena na dan 6. 5. 2026 in ki jo naročnik namerava uporabiti za preverjanje izpolnjevanja izpostavljenih zahtev.

Državna revizijska komisija uvodoma pojasnjuje, da je naročnik pri identifikaciji svojih potreb in oblikovanju predmeta javnega naročila načeloma samostojen, saj je on tisti, ki najbolj pozna svoje delovanje, objektivne potrebe in specifične zahteve, z identifikacijo in opredeljevanjem potreb pa prevzema tudi odgovornost za svoje poslovanje. Določanje in oblikovanje potreb, na podlagi katerih naročnik oblikuje predmet javnega naročila, je tako po naravi stvari v prvi vrsti vezano na odgovornost naročnika za pravilno in zakonito izvajanje njegovih nalog, kot takšno pa predstavlja naročnikovo poslovno odločitev, v katero pravila o javnem naročanju neposredno ne

¹ Kot izhaja iz točke »4. *Pojmovnik*« tehničnih specifikacij, LLM ali LLMs (»*Large Language Models*«) pomeni »*velik(i) jezikovni model(i) za razumevanje/generiranje besedila*«.

morejo posegati (prim. npr. odločitve Državne revizijske komisije v zadevah št. 018-203/2019, 018-020/2020, 018-097/2023, 018-025/2024).

Ne glede na pojasnjeno naročnikovo samostojnost pri oblikovanju predmeta javnega naročila pa mora slednji tehnične zahteve oblikovati skladno z 68. členom ZJN-3, ki med drugim v četrtem odstavku določa, da morajo tehnične specifikacije vsem gospodarskim subjektom zagotavljati enak dostop do postopka javnega naročanja in neupravičeno ne smejo ovirati odpiranja javnih naročil konkurenci. Poleg tega naročnika pri oblikovanju tehničnih zahtev zavezujejo tudi splošna načela javnega naročanja, ki od njega zahtevajo, da jih določi na način, ki zagotavlja konkurenco med ponudniki in njihovo enako obravnavo (5. in 7. člen ZJN-3), ter v obsegu, ki je potreben in sorazmeren z naravo, vsebino, namenom in obsegom predmeta naročila (8. člen ZJN-3).

V tej zvezi gre poudariti še, da načela zagotavljanja konkurence med ponudniki ni mogoče interpretirati v smislu zahteve po vzpostavljanju konkurenčnosti na tistih področjih oziroma v tistih primerih, ko te iz upravičenih razlogov ni mogoče doseči, prav tako pa tudi načela enakopravnosti v pravu javnih naročil ni mogoče razumeti kot absolutne kategorije. Enakopravnost namreč ne pomeni, da mora naročnik v postopku oddaje javnega naročila vsem potencialnim ponudnikom omogočiti enak položaj. Zaradi različnih ekonomskih, tehničnih, kadrovskih in tudi naravnih danosti je dejanski položaj ponudnikov in njihovih ponudb različen, prednosti, ki jih te zagotavljajo, pa je dovoljeno in pogosto celo gospodarno upoštevati. Zato zgolj dejstvo, da naročnik z določeno zahtevo ustvarja razlikovanje med (potencialnimi) ponudniki, še ne pomeni, da je takšna zahteva že sama po sebi diskriminatorna. V naravi same zahteve namreč je, da (potencialne) ponudnike razvršča na tiste, ki določeno zahtevo izpolnjujejo, ter na tiste, ki te zahteve ne izpolnjujejo in posledično ne morejo sodelovati v postopku oddaje javnega naročila. Naročniki so tako v postopkih oddaje javnih naročil upravičeni postavljati zahteve, ki imajo za posledico razlikovanje ponudnikov, vendar le iz razlogov, ki so neposredno povezani in sorazmerni s predmetom javnega naročila ter so objektivno opravičljivi. Razlikovanje ponudnikov glede na kriterije, ki niso objektivno opravičljivi in pomenijo zlasti krajevno, predmetno ali osebno diskriminacijo, s čimer je določen gospodarski subjekt (v primerjavi z drugimi) bodisi postavljen v bistveno slabši položaj bodisi je privilegiran, ne da bi za to obstajali utemeljeni razlogi, pa po drugi strani ni dopustno (v tej zvezi smiselno prim. tudi sodbi Sodišča Evropske unije v zadevah C-513/99 in C-195/21).

V nadaljevanju je tako treba ugotoviti, ali sta izpodbijani zahtevi, tj. uporaba lestvice Arena za ugotavljanje kakovosti ponujenih jezikovnih modelov ter zahtevani točkovni prag za najzmogljivejšo stopnjo na tej lestvici, (v okviru vlagateljevih zatrjevanj) povezani in sorazmerni s predmetom javnega naročila oziroma ali ima naročnik objektivno utemeljene razloge za s strani vlagatelja zatrjevano neupravičeno omejevanje in razlikovanje med ponudniki na podlagi teh zahtev.

Pri tem je treba upoštevati, da naročnik, skladno z ustaljeno prakso Državne revizijske komisije, razlogov, ki opravičujejo določitev posameznih zahtev, ni dolžan navajati (že) v dokumentaciji v zvezi z oddajo javnega naročila, temveč jih je dolžan navesti šele v (morebitnem) postopku pravnega varstva, pa tudi to le v primeru, če (najprej) vlagatelj izpolni t. i. trditveno-dokazno breme in v zahtevku za revizijo zatrjuje ter dokaže dejstva, ki kažejo na to, da naročnik objektivno opravičljivih razlogov za določitev izpodbijanih zahtev ni imel. Takšno stališče je zavzelo tudi Sodišče Evropske unije v zadevi C-568/24, v kateri je v zvezi z vprašanjem objektivne utemeljitve tehničnih specifikacij v razpisni dokumentaciji med drugim pojasnilo, da »*niti člen 18(1) Direktive 2014/24 niti člen 42(1) in (2) te direktive ne nalagata razlage načel transparentnosti in enakega obravnavanja tako, da od javnega naročnika zahtevata, da na datum objave zadevnega obvestila o javnem naročilu navede objektivne utemeljitve, na katerih temeljijo tehnične specifikacije, ki jih je pripravil*« (gl. točko 43 obrazložitve te sodbe, kot tudi ostale točke od 33-48).

Predmet presoje Državne revizijske komisije v danem primeru zato ni vprašanje, ali so bili razlogi za določitev spornih zahtev navedeni že v dokumentaciji v zvezi z oddajo javnega naročila, temveč ali je naročnik v postopku pravnega varstva uspel izkazati njihovo objektivno utemeljenost.

Vlagateljevo izpodbijanje naročnikove zahteve po uporabi lestvice Arena za ugotavljanje zmogljivosti ponujenih jezikovnih modelov temelji na zatrjevanju, da je dodeljevanje točk posameznemu modelu umetne inteligence na predmetni lestvici subjektivno ter da obstajajo druge primerjalne lestvice, ki naj bi zmogljivost jezikovnih modelov odražale bolj objektivno oziroma reprezentativno. Glede na to, da lestvica Arena temelji na subjektivnih preferencah uporabnikov, vlagatelj zatrjuje, da ta ne predstavlja objektivnega kriterija tehnične kakovosti jezikovnih modelov, saj da meri zgolj všečnost oziroma uporabniške preference, ne pa dejanske tehnične ali funkcionalne zmogljivosti modelov, ter da naročnik ni izkazal povezave med rezultatom na tej lestvici ter potrebami slovenskega izobraževalnega in raziskovalnega okolja.

Državna revizijska komisija ugotavlja, da navedeni očitki sami po sebi še ne izkazujejo, da uporaba lestvice Arena glede na predmet javnega naročila ni objektivno utemeljena oziroma da ni primerna za preverjanje izpolnjevanja tehničnih zahtev. Dejstvo, da posamezna primerjalna lestvica temelji na določeni metodologiji ocenjevanja, ki vključuje tudi vrednotenje uporabniških preferenc, namreč (še) ne izkazuje njene neprimernosti ali nezanesljivosti. Naročnik je v postopku pravnega varstva konkretizirano pojasnil, zakaj je izbral prav lestvico Arena. Navedel je, da jo je izbral ravno zato, ker temelji na vrednotenju oziroma primerjavi odgovorov jezikovnih modelov v dejanskih uporabniških interakcijah, pri čemer različica »style-controlled« zmanjšuje vpliv zgolj slogovnih razlik med odgovori, rezultat pa po njegovi oceni predstavlja celovito oceno splošne kakovosti uporabniške izkušnje širokega kroga uporabnikov, katerim je namenjen predmet javnega naročila (tj. učencem, dijakom, študentom, učiteljem, raziskovalcem in drugim uporabnikom). Obenem je pojasnil, zakaj drugih lestvic, zlasti lestvice n8n AI Benchmark (ki jo predlaga vlagatelj), ni štel za primerljivo merilo kakovosti ponujenih modelov, saj naj bi ocenjevale bistveno ožji nabor zmogljivosti oziroma bile namenjene specifičnim primerom uporabe.

Državna revizijska komisija ocenjuje, da so razlogi, s katerimi je naročnik utemeljil uporabo lestvice Arena, glede na predmet naročila razumni in objektivno povezani s cilji, ki jih naročnik s tem javnim naročilom zasleduje. Vlagatelj pa pri tem po oceni Državne revizijske komisije ni uspel izkazati, da ti razlogi niso povezani in sorazmerni s predmetom javnega naročila ter naročnikovimi potrebami. Njegove navedbe namreč izkazujejo predvsem, da obstajajo tudi druge metodologije vrednotenja jezikovnih modelov in da se posamezni modeli na različnih lestvicah uvrščajo različno, ne izkazujejo pa, da lestvica Arena kot taka ne predstavlja ustreznega ali objektivno primernega kriterija za preverjanje splošne zmogljivosti jezikovnih modelov. Naročnik je, kot že navedeno, ustrezno pojasnil, da je predmet naročila zagotovitev jezikovnih modelov za raznoliko skupino uporabnikov in širok spekter primerov uporabe, pri čemer je kot bistveno štel kakovost odgovorov, kot jo zaznavajo končni uporabniki.

Prav tako Državna revizijska komisija ugotavlja, da vlagatelj ni uspel izkazati, da bi moral naročnik dopustiti uporabo drugih primerjalnih lestvic oziroma drugih metodologij ocenjevanja. Okoliščina, da na trgu obstaja več primerjalnih lestvic, ki uporabljajo različne metodologije in posledično modele razvrščajo različno, sama po sebi še ne pomeni, da bi moral naročnik dopustiti uporabo vseh takšnih lestvic oziroma da kot kriterij za ugotavljanje ustreznosti ponujenih modelov ne bi smel določiti ene izmed njih.

Na navedeni zaključek ne vplivajo niti vlagateljeve navedbe o ločenem preverjanju sposobnosti delovanja modelov v slovenskem jeziku. Dejstvo, da je naročnik poleg splošne zmogljivosti modelov (ločeno) preverjal še posamezne dodatne lastnosti, namreč samo po sebi ne pomeni,

da Arena lestvica ni primerna za vrednotenje tistih lastnosti, zaradi katerih jo je naročnik določil kot kriterij za presojo kakovosti ponujenih jezikovnih modelov.

Upoštevajoč vse navedeno Državna revizijska komisija zaključuje, da vlagatelj ni uspel izkazati, da je naročnik ravnal v nasprotju z ZJN-3, ko je za preverjanje zmogljivosti jezikovnih modelov določil uporabo lestvice Arena.

Drugačno pa je vprašanje, ali je naročnik uspel objektivno utemeljiti konkretno določeni minimalni prag 1440 točk na navedeni lestvici. Dejstvo, da je naročnik upravičeno izbral lestvico Arena kot izkaz kakovosti ponujenega jezikovnega modela, namreč še ne pomeni, da je hkrati izkazal tudi objektivno utemeljenost konkretno določene mejne vrednosti, ki jo morajo modeli na tej lestvici dosegati.

Naročnik sicer obširno pojasnjuje metodologijo, po kateri je modele razvrstil v tri kakovostne razrede, pri čemer se sklicuje na sodelovanje z akademsko skupnostjo ter na analogijo z didaktičnimi teorijami določanja minimalnih standardov za preverjanje znanja. Na tej podlagi je določil razrede najboljših (odličnih/vrhunskih) okvirno 15 % modelov, dobrih naslednjih približno 25 % modelov in zadostnih (še sprejemljivih) naslednjih okvirno tretjino vseh modelov, uvrščenih na lestvici Arena. Vendar navedena delitev, ki jo je naročnik sam ustvaril, (še) ne izkazuje, zakaj prav meja 1440 točk predstavlja objektivno potrebno tehnično zahtevo za izvedbo predmetnega naročila. Naročnik namreč ni pojasnil, na podlagi katerih konkretnih strokovnih analiz, testiranj ali drugih preverljivih okoliščin je ugotovil, da modeli, uvrščeni tik pod navedeni prag, njegovih potreb ne bi več mogli zadovoljiti, oziroma da razlika med modeli nad in pod pragom pomeni takšno razliko v zmogljivosti, ki bi bila za izvedbo predmeta naročila bistvena. Prav tako ni pojasnil, zakaj naj bi bila prav uvrstitev med približno najboljših 15 % modelov nujna za izvajanje nalog, ki jih je opredelil kot najzahtevnejše.

Enako velja tudi za navedbe o številu modelov, ki dosegajo zahtevani prag. Dejstvo, da prag izpolnjuje večje število modelov različnih ponudnikov, sicer lahko kaže, da zahteva sama po sebi ni oblikovana diskriminatorno, ne izkazuje pa še, da je prav meja 1440 točk objektivno potrebna in sorazmerna. Prav tako naročnik ni uspel izkazati, da bi nižji prag nujno omogočil ponudbo modelov, ki za njegove potrebe ne bi bili več ustrezni. Njegove navedbe, da bi ponudniki sicer lahko ponudili cenejše modele nižje kakovosti, ostajajo na ravni splošnih trditev, saj niso podprte z objektivnimi podatki, iz katerih bi izhajalo, da modeli pod pragom 1440 točk dejansko ne dosegajo zmogljivosti, potrebnih za izvedbo predmeta naročila.

Državna revizijska komisija zato ugotavlja, da naročnik ni uspel izkazati utemeljenih razlogov za določitev minimalnega praga 1440 točk na lestvici Arena, s tem pa je kršil četrti odstavek 68. člena ZJN-3 ter načelo zagotavljanja konkurence med ponudniki (5. člen ZJN-3), načelo enakopravne obravnave ponudnikov (7. člen ZJN-3) in načelo sorazmernosti (8. člen ZJN-3).

V posledici vsega ugotovljenega Državna revizijska komisija zaključuje, da je vlagatelj uspel izkazati, da je naročnik določbo dokumentacije v zvezi z oddajo javnega naročila, ki se nanaša na zahtevo v zvezi s točkovnim pragom za zahtevano stopnjo 1 (tj. najzmogljivejši jezikovni model) na lestvici Arena, oblikoval v nasprotju z določbami ZJN-3.

Glede na navedeno je Državna revizijska komisija zahtevku za revizijo na podlagi druge alineje prvega odstavka 39. člena ZPVPJN ugodila tako, da je delno razveljavila postopek javnega naročanja na način, da je v dokumentaciji v zvezi z oddajo javnega naročila za tretjo fazo v tehničnih specifikacijah razveljavila zahtevo, določeno v četrti alineji točke 5.1.2, ki se glasi: »[M] jezikovni model (za zahtevano stopnjo 1 - najzmogljivejši) dosega vsaj ≥ 1440 točk na Leaderboard Overview <https://arena.ai/leaderboard-> (Text Arena - Overall (style-controlled))«.

Skladno s tretjim odstavkom 39. člena ZPVPJN mora dati Državna revizijska komisija naročniku napotke za pravilno izvedbo postopka v delu, ki je bil razveljavljen. Ker je ugotovljene kršitve mogoče odpraviti le na stopnji postopka oddaje javnega naročila pred potekom roka za predložitev ponudb, Državna revizijska komisija naročnika napotuje, da v primeru, če se bo odločil za nadaljevanje postopka javnega naročanja, dokumentacijo v zvezi z oddajo javnega naročila ob upoštevanju ugotovitev Državne revizijske komisije iz tega sklepa ustrezno spremeni in določi nov rok za predložitev ponudb ter njihovo odpiranje.

S tem je utemeljena odločitev Državne revizijske komisije iz 1. točke izreka tega sklepa.

Vlagatelj v zahtevku za revizijo zahteva povrnitev stroškov, nastalih v postopku pravnega varstva, in sicer povračilo vplačane takse v višini 4.000,00 EUR.

Če je zahtevek za revizijo utemeljen, mora naročnik iz lastnih sredstev vlagatelju povrniti potrebne stroške, nastale v predrevizijskem in revizijskem postopku, vključno s takso (tretji odstavek 70. člena ZPVPJN).

Če bi vlagatelj v celoti uspel z zahtevkom za revizijo, bi mu Državna revizijska komisija, upoštevajoč prvi, drugi in tretji odstavek 70. člena ZPVPJN, kot potrebne priznala celotne priglašene stroške za dolžno vplačano takso v višini 4.000,00 EUR. Ker pa je vlagatelj z zahtevkom za revizijo uspel le v eni polovici (tj. v delu, ki se nanaša na zahtevo glede točkovnega praga na lestvici Arena, ne pa tudi v delu, ki se nanaša na uporabo te lestvice), mu je Državna revizijska komisija priznala stroške v višini ene polovice priglašenih stroškov.

Državna revizijska komisija je tako vlagatelju priznala stroške v višini 2.000,00 EUR, ki mu jih je naročnik dolžan povrniti v roku 15 dni od prejema tega sklepa, skupaj z zakonskimi zamudnimi obrestmi, ki tečejo od prvega dne po poteku roka za prostovoljno izpolnitev do plačila.

S tem je utemeljena odločitev Državne revizijske komisije iz 2. točke izreka tega sklepa.

Naročnik je v odločitvi o zahtevku za revizijo odločil tudi, da mu mora vlagatelj povrniti stroške postopka, ki so mu nastali zaradi odločanja o zahtevku za revizijo. Naročnik je navedel, da je pri pripravi sklepa sodeloval en uslužbenec, katerega stroški izhajajo iz naslednje opredelitve: strošek plače javnega uslužbenca na delovnem mestu podsekretar na mesečnem nivoju znaša 2.818,77 EUR, kar pomeni, da vrednost ene ure, ob upoštevanju 176 ur delovne obveznosti, znaša 16,02 EUR. Ker je bilo za pripravo predmetne odločitve porabljenih 8 ur, naročnik ocenjuje strošek v višini 128,16 EUR.

Naročnikovo stroškovno zahtevo je treba zavrniti že iz razloga, ker naročnik v predrevizijskem postopku na podlagi določb ZPVPJN sprejema odločitve v vlogi organa odločanja oziroma pravnega varstva (1. alineja 2. člena ZPVPJN v povezavi z 28. členom ZPVPJN), ne pa v vlogi stranke predrevizijskega postopka (prvi odstavek 3. člena ZPVPJN). Zato naročnik do povrnitve priglašenih stroškov ni upravičen.

S tem je utemeljena odločitev Državne revizijske komisije iz 3. točke izreka tega sklepa.

Pravni pouk:

Upravni spor zoper to odločitev ni dovoljen.

Predsednik senata:

Aleksander Petrovčič, univ. dipl. ekon.,
član Državne revizijske komisije

Vročiti:

- vlagatelju,
- naročniku,
- ministrstvu, pristojnemu za javna naročila.

Vložiti:

- v spis zadeve.