

U 23/2024 rd

Statsrådets skrivelse till riksdagen om Finlands anslutning till det europeiska konsortiet för digital infrastruktur för alliansen för språkteknologi

I enlighet med 96 § 2 mom. i grundlagen översänds till riksdagen en promemoria om Finlands anslutning till det europeiska konsortiet för digital infrastruktur för alliansen för språkteknologi samt konsortiets stadgar.

Helsingfors den 8 Maj 2024

Kommunikationsminister Lulu Ranne

Specialsakkunnig Tomi Paavola

26.3.2024

FÖRSLAG OM FINLANDS ANSLUTNING TILL DET EUROPEISKA KONSORTIET FÖR DIGITAL INFRASTRUKTUR FÖR ALLIANSEN FÖR SPRÅKTEKNOLOGI

1 Finlands anslutning till det europeiska konsortiet för digital infrastruktur för alliansen för språkteknologi

Finland har för avsikt att ansluta sig som observatör till det europeiska konsortiet för digital infrastruktur för alliansen för språkteknologi (*Alliance for Language Technologies EDIC*, senare ALT-EDIC). ALT-EDIC utvecklar en gemensam europeisk språkinfrastruktur särskilt för stora språkmodeller. Målet är att stärka Europas konkurrenskraft, förbättra tillgången till europeiska språkdata samt upprätthålla Europas språkliga mångfald och kulturella rikedom. Finlands mål är att senare ansluta sig som ordinarie medlem till ALT-EDIC, om behövlig finansiering för de medlemsavgifter som medlemskapet förutsätter säkerställs.

EU offentliggjorde som ett led i genomförandet av målen i policyprogrammet för det digitala decenniet 2030 europeiska konsortier för digital infrastruktur (*European Digital Infrastructure Consortium*, senare EDIC) som en form av juridisk personlighet med vars hjälp inrättandet och genomförandet av europeiska flerlandsprojekt kan påskyndas och förenklas. Genom flerlandsprojekt underlättas uppnåendet av EU:s digitala mål.

Varje EDIC är en juridisk person som inrättas genom kommissionens beslut på ansökan av minst tre EU-medlemsstater. De inrättande medlemsstaterna fastställer EDICs förvaltningsstruktur och andra bestämmelser om verksamheten i konsortiets stadgar. EDICs budget baserar sig på medlemmarnas finansiella bidrag, som kompletteras med andra inkomstkällor, såsom EU:s eller nationella bidrag. EDIC kan genomföra ett flerlandsprojekt genom att enligt de inrättande medlemsstaternas prövning ta i bruk gemensam infrastruktur, tillhandahålla tjänster och sammanföra företrädare för offentliga organ, privata aktörer, slutanvändare och olika sektorer.

ALT-EDIC har inrättats genom kommissionens beslut av den 1 februari 2024 ((EU) 2024/458). Den 5 december 2023 uppmanade Bulgarien, Frankrike, Irland, Italien, Kroatien, Lettland, Litauen, Nederländerna, Polen och Slovenien kommissionen att inrätta ALT-EDIC. Grekland anslöt sig till denna begäran den 16 januari 2024. Frankrike är värdmedlemsstat för ALT-EDIC och dess stadgeenliga säte är i Villers-Cotterêts. ALT-EDIC har inrättats för att genomföra ett flerlandsprojekt i syfte att utveckla en gemensam europeisk infrastruktur för språkteknologi.

ALT-EDICs medlemmar med observatörsstatus behöver inte betala medlemsavgift för att höra till konsortiet. Observatörer och andra parter som inte är ordinarie medlemmar i ALT-EDIC kan göra frivilliga donationer till ALT-EDIC som inte ska betraktas som medlemsavgift. Senare, om Finland eventuellt blir ordinarie medlem i ALT-EDIC, består Finlands kostnader av den stadgeenliga medlemsavgiften, som skulle vara 79 268,26 euro per år. Det är möjligt att erlägga finansiella bidrag i anslutning till medlemsavgiften kontant eller in natura.

Medlemmen har rätt att delta och rösta i ALT-EDICs generalförsamling och delta i andra evenemang som den ordnar samt att få det stöd som ALT-EDIC tillhandahåller för att utveckla

relevanta system, processer och tjänster. Medlemmen har också rätt att delta i projektförslag och finansieringsmöjligheter på europeisk och nationell nivå, där ALT-EDIC fungerar som samordnare, medlem i konsortiet eller förmånstagare. En observationsmedlem har rätt att delta i dess stämma utan rösträtt och i evenemang som den ordnar.

De medlemmar som ansluter sig till konsortiet förbinder sig att höra till och betala medlemsavgift till konsortiet de första tre åren efter inrättandet av ALT-EDIC. Observatörer behöver inte förbinda sig till konsortiet för tre år efter inrättandet av konsortiet.

Varje medlem och observatör ska anmäla sin företrädare till generalförsamlingen. Enligt ALT-EDICs stadgar ska Finland representeras av kommunikationsministeriet, som utser företrädare till ALT-EDICs generalförsamling.

2 Finlands deltagande i ALT-EDIC

Finlands mål är att främja språklig och kulturell mångfald i Europa genom att stödja särskilt finskans och svenskans samt de i Finland talade minoritetsspråkens ställning vid utvecklingen av språkteknologier. Finland stödjer också teknologisk spetskompetens, utformningen av en dataarkitektur samt förklarbara och återanvändningsbara språkmodeller.

Finland strävar efter att främja de åtgärder som anges i ALT-EDICs stadgar som följer:

1. Data: Tillgången på språkdata är fortfarande en utmaning. Finland kan delta i insamlingen av högklassigt material samt material som företräder en bredare språkanvändning genom inkluderande och representativ datainsamling för språkmodeller. Detta har testats och dokumenteras bl.a. på den prisbelönta webbplatsen lahjoitapuhetta.fi, där man har samlat tusentals timmar kollokvialt tal för både akademiskt och industriellt bruk.
2. Befintliga språkmodeller: Finlands akademiska intressentgrupper möjliggör återanvändning av befintliga språkmodeller och utvecklar nya modeller med öppen källkod för industriella aktörers återanvändning samt erbjuder metoder att utveckla flerspråkiga modeller mellan finska och svenska samt andra språk. I olika språkteknologiska projekt utvecklas särskilda metoder för att finjustera, mäta och optimera befintliga modeller och tillgodose särskilda behov samt för att förbättra små och medelstora företags och europeiska företags närhet.
3. Modellutveckling: Finland erbjuder en plats för ett europeiskt högpresterande datorsystem i superdatormiljön Lumi EuroHPC.
4. Utvärdering, certifiering och normalisering: Utvärdering, certifiering och normalisering av nya språkmodeller är av väsentlig vikt för att inge förtroende hos människor samt för att iakttä nationell och EU:s lagstiftning. I Finland finns kompetens att tillhandahålla metoder som fokuserar särskilt på behandlingsmodeller för naturligt språk (*Natural Language Processing*, senare NLP-modell), som upptäcker och utvärderar bl.a. hallucinationer, eventuell diskriminering och snedvridningar som kan hänföra sig till exempelvis kön, accent, ålder eller utbildningsnivå.
5. Ekosystem: Projektet LAREINA (*Language Resource Infrastructure for AI*) stödjer ett ekosystem som bildats av 50 finländska organisationer. I projektet utvecklas och främjas kontakterna mellan industrin, forskningen och den offentliga sektorn i Finland genom att uppmuntra teknologisk överförbarhet och spridning av forskningsresultat.

Projektet stödjer medlemsinstitutionerna särskilt vid införandet av språkmodeller. Det visar vägen för Finlands genomförande av ALT-EDIC genom att fungera som rådgivare och genomföra EU:s datapolicy.

3 Det europeiska konsortiet för digital infrastruktur för alliansen för språkteknologi

3.1 Konsortiets uppgifter

ALT-EDICs uppgift är att genomföra ett flerlandsprojekt i syfte att utveckla en gemensam europeisk infrastruktur för språkteknologi. Genom projektet ökas Europas förmåga att förbättra konkurrenskraften med beaktande av små och medelstora företags ställning och upprätthålls Europas språkliga mångfald och kulturella rikedom. ALT-EDICs syfte är att åtgärda bristen på europeisk språkdata som kan användas för utbildning om AI-lösningar. ALT-EDIC främjar skapandet av språkmodeller för officiella och regionala språk i Europa.

För att genomföra konsortiets uppgift och mål har ALT-EDIC ett åtgärdsprogram. Åtgärderna samordnas med andra europeiska initiativ, såsom den gemensamma europeiska forskningsinfrastrukturen för språkresurser och språkteknologi (*Common Language Resources and Technology Infrastructure*, senare CLARIN-ERIC). Finland är medlem i CLARIN-ERIC. Det är en forskningsinfrastruktur som främjar humanistisk och samhällsvetenskaplig forskning genom att tillhandahålla forskare en forskningsinfrastruktur som möjliggör tillgång till språkresurser och verktyg på europeisk nivå. Forskningsinfrastrukturen består av de deltagande ländernas nationella forskningsinfrastrukturer, vilkas kompabilitet och sambruk CLARIN-ERIC koordinerar. Den erbjuder också möjligheter att utveckla ny språkteknologi och tillämpningar. Om Finland senare ansluter sig till ALT-EDIC när den finansiering som behövs har blivit en egentlig medlem, är det väsentligt att främja en förvaltningsövergripande tydlig rollfördelning och ett nära samarbete mellan ALT-EDIC och CLARIN-ERIC i syfte att utveckla tjänsterna och eventuellt i finansieringen. På så sätt kan man identifiera synergier mellan de två organisationernas verksamhet.

3.2 Medlemmar och observatörer samt deras rättigheter och skyldigheter

Europeiska ekonomiska samarbetsområdets medlemsstater och EES-staternas administrativa regioner kan bli medlemmar i ALT-EDIC. Det är också möjligt att ansluta sig till konsortiet som observatör. Observatörer har rätt att delta i generalförsamlingen utan rösträtt. Tredjeländer och internationella organisationer kan endast bli observatörer. Beslut om godkännande av nya medlemmar fattas av generalförsamlingen på ansökan i enlighet med det förfarande som anges i stadgarna.

Medlemmarna har rätt att utse den institution som företräder dem i generalförsamlingen och att delta i ALT-EDICs generalförsamlings verksamhet och utöva sin rösträtt. En medlemsstat i Europeiska ekonomiska samarbetsområdet som är medlem eller observatör kan företräddas av en eller flera offentliga aktörer eller en aktör från en administrativ region eller en privat aktör som sköter en offentlig serviceuppgift. Enligt stadgan ska medlemmarna betala en årlig medlemsavgift. Grunderna för hur den fastställs beskrivs i bilaga III till stadgan. Varje medlem och observatör är skyldig att främja ALT-EDICs verksamhet i enlighet med stadgarna.

En medlem får lämna ALT-EDIC eller uteslutas i enlighet med det förfarande som anges i stadgarna. Under de första tre åren efter inrättandet av ALT-EDIC får en medlem inte lämna konsortiet, om man inte har överenskommit om ett kortare medlemskap i samband med anslutningen. Observatörer behöver inte förbinda sig till konsortiet tre år efter inrättandet av

konsortiet och kan lämna konsortiet genom en skriftlig anmälan till generalförsamlingens ordförande.

3.3 Förvaltning och beslutsfattande

ALT-EDICs förvaltningsstruktur omfattar generalförsamlingen, generaldirektören, en strategisk styrkommitté (*Strategic Orientation Committee*), en vetenskaplig och teknisk expertgrupp samt en juridisk och etisk expertgrupp. Kommittén stödjer generalförsamlingens och generaldirektörens verksamhet. Expertgrupperna stödjer generalförsamlingens verksamhet med sin expertis.

ALT-EDICs högsta beslutanderätt utövas av generalförsamlingen, som består av företrädare för medlemmarna och observatörerna. Den utser och beviljar avsked åt konsortiets generaldirektör samt godkänner nya medlemmar, genomförandebestämmelser och genomförandestrategin, den årliga budgeten, arbetsprogrammet, den fleråriga finansieringsplanen och bokslutet. Den beslutar om medlemmarnas och observatörernas status och finansiella bidrag. Den inrättar vid behov rådgivande organ och utser medlemmar till dem. Generalförsamlingen kan ändra ALT-EDICs stadgar.

Varje medlem eller observatör utser en officiell företrädare till generalförsamlingen. Tillsammans med varje medlems eller observatörs företrädare kan dessutom experter delta i generalförsamlingen så att högst fem personer får ingå i delegationen. I generalförsamlingen har varje medlem rösträtt i förhållande till dess finansiella bidrag, som bestäms i enlighet med bilaga II till stadgarna:

- årligt finansiellt bidrag från 50 000 euro till 249 999 euro: 1 röst
- årligt finansiellt bidrag från 250 000 euro till 499 999 euro: 2 röster
- årligt finansiellt bidrag minst 500 000 euro: 3 röster

Observatörer har inte rösträtt i generalförsamlingen.

3.4 Finansiering

ALT-EDICs verksamhet finansieras med medlemsavgifter och genom att ansöka om finansiering från bl.a. EU:s olika finansiella instrument. Medlemsavgifterna bestäms i enlighet med principer som fastställts i stadgarna och bilaga III till dem. Det är möjligt att erlägga finansiella bidrag som hänför sig till medlemsavgiften kontant och in natura. Observatörer och andra parter som inte är ordinarie medlemmar i ALT-EDIC kan göra frivilliga donationer till ALT-EDIC. Donationer ska inte betraktas som medlemsavgift.

Om Finland blir ordinarie medlem i ALT-EDIC, skulle Finlands medlemsavgift vara 79 268,26 euro per år under perioden 2024-2027. Den första medlemsavgiften täcker åren 2024 och 2025.

I enlighet med beslutet om inrättande av ALT-EDIC ((EU) 2024/458) är konsortiet och dess medlemmar undantagna från mervärdesskatt och punktskatt. Undantagen från mervärdesskatt begränsas till ALT-EDICs och medlemmarnas inköp avsedda att användas officiellt och uteslutande av konsortiet. Förutsättningen för undantag är att dessa inköp endast görs för ALT-EDICs icke-ekonomiska verksamhet i överensstämmelse med dess verksamhetsprinciper. Undantag från mervärdesskatt begränsas till inköp vars värde överstiger 300 euro. Undantag från punktskatt begränsas till ALT-EDICs inköp avsedda att användas officiellt och uteslutande av ALT-EDIC. Förutsättningen för undantag är att i) dessa inköp endast görs för ALT-EDICs

icke-ekonomiska verksamhet i överensstämmelse med dess verksamhetsprinciper och ii) värdet på inköpet överstiger 300 euro.

4 Bakgrund

I ALT-EDICs stadgar avses med språkteknologi (*Language Technologies*, senare LT-teknologi) digitala metoder för att analysera, producera eller ändra mänskligt språk i alla dessa former och därtill anknytande betydelser även tillsammans med andra former av kommunikation, såsom video, bilder, musik osv. LT-teknologi kan utnyttjas bl.a. vid automatisering av olika arbetsuppgifter och processer. I Finland har språkteknologisk forskning och utveckling redan bedrivits under flera årtionden.

ALT-EDIC har koppling till utveckling av den europeiska språkdatarymden (*The Common European Language Data Space*, senare LDS). LDS är en av EU:s många datarymder och den skapar en förvaltningsstruktur för informationsutbyte mellan olika sektorer. Med hjälp av LDS kan man utveckla tjänster och verktyg som baserar sig på språkteknologier. Även ALT-EDIC har tillgång till data som utnyttjas i LDS.

Finland har varit med och inrättat och anslutit sig till flera olika EDIC, med teman som bl.a. mobilitets- och logistikdata, blockkedjeteknik, digital tvilling av den byggda miljön, cybersäkerhet, genomik, offentlig förvaltning och livsmedelsproduktion.

Anslutningen till ALT-EDIC genomför de mål i statsminister Petteri Orpos regeringsprogram som gäller dataekonomi och digitalisering. När digitalisering och teknologisk utveckling utnyttjas rätt kan de avsevärt förbättra människornas livskvalitet, företags konkurrenskraft och tillgången på arbetskraft samt effektivisera den offentliga förvaltningen. Målet är att fullt ut gripa ta om den potential som nya teknologier och digitaliseringen erbjuder.

5 Förslagets rättsliga grund och beslut om Finlands deltagande

Inrättandet av EDIC och dess funktionssätt styrs av unionsrätten och särskilt beslutet om inrättande av policyprogrammet för det digitala decenniet 2030 ((EU) 2022/2481). Sekundärt tillämpas lagstiftningen i den medlemsstat där EDIC har sitt stadgeenliga säte när det gäller frågor som inte, eller endast delvis, regleras av unionsrätten, särskilt beslut ((EU) 2022/2481). I tredje hand tillämpas stadgarna och genomförandebestämmelser för dessa.

EDIC inrättas med stöd av beslut ((EU) 2022/2481) i ett förfarande där kommissionen efter att ha undersökt om medlemsstaternas ansökan, stadgarna, den tekniska beskrivningen av det europeiska flerlandsprojektet samt värmedlemsstatens förklaring om huruvida den erkänner EDIC som ett internationellt organ som avses i artiklarna från och med dagen för inrättande av EDIC fattar beslut om inrättande av EDIC eller om avslag på ansökan. EDIC är juridiska personer från och med den dag då kommissionens beslut träder i kraft. ALT-EDIC har inrättats den 1 februari 2024.

Enligt 93 § 2 mom. i grundlagen svarar statsrådet för den nationella beredningen av beslut som fattas i Europeiska unionen och beslutar om Finlands åtgärder som hänför sig till dem, om inte beslutet kräver godkännande av riksdagen. Riksdagens grundlagsutskott har ansett att 93 § 2 mom. också täcker frågor som formellt sett inte omfattas av unionens behörighet men som både till sitt innehåll och sina konsekvenser kan jämföras med unionsfrågor (GrUB 10/1998 rd). Enligt 96 § i grundlagen ska statsrådet sedan det fått kännedom om förslag till sådana rättsakter, fördrag eller andra åtgärder om vilka beslut fattas inom Europeiska unionen eller som annars

enligt grundlagen skulle falla inom riksdagens behörighet utan dröjsmål sända förslaget till talmannen genom en skrivelse.

Med stöd av vad som anförts ovan begär kommunikationsministeriet fullmakt av statsrådet att för Finlands del underteckna en ansökan om observatörsstatus i ALT-EDIC och senare en eventuell ansökan om medlemskap. Riksdagen deltar i beredningen i enlighet med 96 § i grundlagen.

6 Förslagets konsekvenser

6.1 Konsekvenser för teknologi och vetenskap

Finlands deltagande i ALT-EDIC främjar vetenskaplig forskning och teknologiska innovationer genom att erbjuda direkt tillträde till utvecklingsarbete och nya teknologier inom en gemensam europeisk infrastruktur för språkteknologi. Detta är viktigt även med tanke på Finlands internationella konkurrenskraft. ALT-EDIC torde erbjuda möjligheter att utveckla nya teknologier och metoder i samarbete med företag. Dessutom förbättrar det möjligheterna till internationellt forskningssamarbete och ansökningar om finansiering på EU-nivå. Konsortiet ökar tillgången på europeisk språkdata. Medlemmar och observatörer i ALT-EDIC har rätt att delta i evenemang och verkstäder som ordnas inom ramen för ALT-EDIC, och som kan stärka den europeiska kompetensen i anslutning till språkteknologier. Konsekvenserna för immateriella rättigheter, såsom upphovsrätt och industriell äganderätt, måste också beaktas i ALT-EDICs verksamhet.

6.2 Samhälleliga konsekvenser

Den gemensamma europeiska språkinfrastruktur som utvecklas inom ALT-EDIC kan utnyttjas för att utveckla innovationer och tjänster som kan svara på både globala utmaningar, såsom klimatförändringen, och lokala frågor som t.ex. utveckling av olika tolkningstekniker och tolktjänster för olika nationella språk och minoritetsspråk eller personer med funktionsnedsättning.

6.3 Ekonomiska konsekvenser

När Finland senare eventuellt ansöker om att bli ordinarie medlem i ALT-EDIC i enlighet med stadgarna ska Finland betala en medlemsavgift på 79 268,26 euro per år under perioden 2024-2027. Det är möjligt att erlagga finansiella bidrag som hänför sig till medlemsavgiften kontant och in natura. Observatörer och andra parter som inte är ordinarie medlemmar i ALT-EDIC kan göra frivilliga donationer till ALT-EDIC. Donationer ska inte betraktas som medlemsavgift.

Generativa AI-teknologier som utvecklats för att producera och hantera språk, dvs. teknologier för stora språkmodeller (*Large Language Model*, senare LLM-teknologi) är förenade med enorm ekonomisk potential. Enligt *McKinsey & Companys* utredning kan generativ AI stå för en andel på upp till 13 miljarder euro av Finlands bruttonationalprodukt 2045 och leda till en produktivitetstillväxt på 0,1–0,6 procentenheter fram till år 2040.¹ LLM-teknologi kan påskynda uppkomsten av nya slags tjänster och innovationer samt möjliggöra till exempel automatisering

¹ *McKinsey & Company*, Generatiivisen tekoälyn taloudellinen potentiaali Suomelle (ung. Generativ AI erbjuder Finland ekonomisk potential):

<https://www.mckinsey.com/fi/-/media/mckinsey/locations/europe%20and%20middle%20east/finland/news/the%20economic%20potential%20of%20generative%20ai%20for%20finland.pdf>

av olika arbetsuppgifter och processer inom såväl den offentliga som den privata sektorn, som torde öka arbetsproduktiviteten.

7 Ålands behörighet

En anslutning till det europeiska konsortiet för digital infrastruktur omfattas inte av Ålands självstyrelse.

8 Den nationella behandlingen av förslaget

Skriftligt förfarande i sektionen för konkurrenskraft (EU8) 8–11.4.2024.

Skriftligt förfarande i sektionen för kommunikation (EU19) 8–11.4.2024.

Skriftligt förfarande i sektionen för forskning och innovationer (EU20) 8–11.4.2024.

9 Statsrådets ståndpunkt

Statsrådet ställer sig positivt till medlemsstaternas samarbete i anslutning till det europeiska konsortiet för digital infrastruktur för alliansen för språkteknologi. Finland är en föregångare inom forskning och utveckling kring stora språkmodeller samt inom högpresterande datorsystem, vilket ger goda utgångspunkter i den internationella konkurrensen. Statsrådet betonar att genom nya teknologier, såsom kvantdatorteknik, robotik och effektivt utnyttjande av AI-lösningar samt högre automatiseringsgrad, skapas ny affärsverksamhet, stöds företagens konkurrenskraft och underlättas arbetskraftsbristen. En anslutning till ALT-EDIC ligger i Finlands intresse.

Statsrådet föreslår att Finland ansluter sig till ALT-EDIC som observatör. Dessutom föreslår statsrådet att Finland senare blir ordinarie medlem i ALT-EDIC, om behövlig finansiering för att betala medlemsavgifterna säkerställs.