ЯРИЛЦЛАГА, НИЙТЛЭЛ · 2026 оны 5-р сарын 11

ХИЙМЭЛ ОЮУНЫ ЭРСДЭЛИЙГ ҮНЭЛЭХ ЭРХ ЗҮЙН МЕХАНИЗМ БА ХҮНИЙ ЭРХИЙН ХАМГААЛАЛТ



УЕПГ-ын дэргэдэх Сургалт, судалгааны Хүрээлэнгийн Сургалтын албаны дарга, доктор, дэд профессор Б.Өнөрмаагийн "Хиймэл оюуны эрсдэлийг үнэлэх эрх зүйн механизм ба хүний эрхийг хамгаалалт" илтгэлийг хүргэж байна.

Нэг. Оршил

Хиймэл оюун ухаан (Künstliche Intelligenz, цаашид “AI” гэх) нь сүүлийн жилүүдэд нийгэм, эдийн засаг, төрийн үйл ажиллагаанд өргөн хүрээнд нэвтэрч, мэдээлэл боловсруулах, дүн шинжилгээ хийх, автомат шийдвэр гаргах үйл явцыг шинэ түвшинд хүргэж байна[1]. AI-д суурилсан системүүд нь үр ашиг, хурд, нарийвчлалыг нэмэгдүүлж байгаа боловч тэдгээрийн хэрэглээ нь хүний үндсэн эрх, эрх чөлөөнд сөргөөр нөлөөлөх шинэ төрлийн эрсдэлийг үүсгэх болсон[2].

Ялангуяа AI системийг төрийн байгууллага, хууль сахиулах болон шүүхийн ажиллагаанд ашиглах явцад хүний эрхийн зөрчил үүсэх эрсдэл нэмэгдэж байгаа нь олон улсын түвшинд онцгой анхаарал татаж байна.

Тухайлбал, алгоритмын шийдвэр гаргалт, өгөгдөлд суурилсан профайлчлал мөн ил тод бус ажиллагаа буюу “хар хайрцаг”-ийн шинж — өөрөөр хэлбэл AI системийн

шийдвэр гаргах дотоод логик, ашигласан өгөгдөл, шийдвэрт хүрсэн үндэслэл нь хэрэглэгч болон хяналт хэрэгжүүлэгч байгууллагад ойлгомжгүй, тайлбарлагдах боломж хязгаарлагдмал байх нөхцөл — нь хүний эрхийн баталгааг хангах уламжлалт эрх зүйн механизмд бодит сорилт учруулж байна[3].

Энэхүү нөхцөл байдал нь AI-ийн хэрэглээг зөвхөн технологийн дэвшлийн асуудал бус, харин хүний эрхийн хамгаалалт, эрх зүйн зохицуулалтын хүрээнд цогцоор авч үзэх шаардлагыг бий болгож байна. Үүнтэй холбогдуулан олон улсын болон бүс нутгийн түвшинд AI-ийн эрсдэлийг үнэлэх, ангилах, бууруулахад чиглэсэн эрх зүйн зохицуулалтыг боловсруулж эхэлсний нэг жишээ нь Хиймэл оюуны тухай Европын холбооны журам (“EU Artificial Intelligence Act”) юм[4].

Тус журам нь AI системийг эрсдэлийн түвшнээр ангилан, ялангуяа өндөр эрсдэлтэй (Hochrisiko-KI) системд тусгай шаардлага тогтоох замаар хүний эрхийг хамгаалах зорилготой “эрсдэлд суурилсан хандлага” (risikobasierter Ansatz)-ыг нэвтрүүлсэн нь онцлог юм[5].

Хоёр. AI-ийн эрсдэлийн үнэлгээний эрх зүйн механизм

AI-ийн хэрэглээ өргөжин тэлж, хүний эрхэд үзүүлэх нөлөө нь нэмэгдэж буй нөхцөлд түүний эрсдэлийг урьдчилан тодорхойлох, үнэлэх, ангилах эрх зүйн механизм бий болгох шаардлага үүссэн байна[6]. Энэ хүрээнд Европын холбооноос AI-ийн зохицуулалтад “эрсдэлд суурилсан хандлага”-ыг нэвтрүүлсэн нь онцлог бөгөөд уг хандлага нь AI системийг хүний эрх, аюулгүй байдалд үзүүлэх эрсдэлийн түвшнээр нь ялган зохицуулахад чиглэж байна[7].

Хиймэл оюуны тухай Европын холбооны журам (“EU Artificial Intelligence Act”) нь AI системийг дараах дөрвөн ангилалд хуваан үздэг. Үүнд:

1. Хүлээн зөвшөөрөх боломжгүй эрсдэлтэй (unannehmbares Risiko) AI систем[8]

Энэ ангилалд хүний эрх, нэр төр, эрх чөлөөнд ноцтой зөрчил учруулах шинжтэй AI системүүд хамаарах бөгөөд ийм төрлийн хэрэглээг бүрэн хориглосон байна. Тухайлбал, хүний зан үйлд нөлөөлөх замаар далд аргаар шийдвэр гаргалтад нөлөөлөх систем, эсхүл эмзэг бүлгийг ашиглан сэтгэл зүйн дарамт үзүүлэх AI хэрэглээ нь энэ ангилалд багтана.

2. Өндөр эрсдэлтэй (Hochrisiko-KI) систем[9]

Энэ нь хүний эрх, аюулгүй байдалд ноцтой нөлөө үзүүлэх боломжтой AI системүүдийг хамрах бөгөөд төрийн удирдлага, хууль сахиулах, шүүхийн ажиллагаа, боловсрол, хөдөлмөр эрхлэлт зэрэг салбарт ашиглагдах AI системүүд ихэвчлэн энэ ангилалд хамаарна. Ийм системд дараах үндсэн шаардлагыг тавьдаг. Үүнд:

Эрсдэлийн удирдлагын тогтолцоог нэвтрүүлэх

Өгөгдлийн чанар, найдвартай байдлыг хангах

Ил тод байдлыг хангах

Хүний хяналт хэрэгжүүлэх

Баримтжуулалт, хяналтын механизмыг бүрдүүлэх

Эдгээр шаардлага нь AI системийн үйл ажиллагааг хүний эрхэд нийцүүлэн хязгаарлах, хяналт тавих эрх зүйн баталгааг бүрдүүлж байна.

3. Хязгаарлагдмал эрсдэлтэй (begrenztes Risiko) систем[10]

Энэ ангилалд багтах AI системүүд нь хүний эрхэд ноцтой нөлөө үзүүлэхгүй боловч тодорхой мэдээллийн ил тод байдлыг хангах шаардлагатай байдаг. Жишээлбэл, хэрэглэгч AI-тэй харилцаж байгаагаа мэдэх боломжтой байх зэрэг шаардлага тавигддаг.

4. Хамгийн бага эрсдэлтэй (minimales Risiko) систем[11]

Энэ төрлийн AI системүүд нь хүний эрхэд үл мэдэгдэх буюу маш бага нөлөө үзүүлэх тул тусгай эрх зүйн хатуу зохицуулалт шаардахгүй бөгөөд сайн дурын зарчимд тулгуурласан хэрэглээг зөвшөөрдөг байна.

Ийнхүү AI-ийн эрсдэлийг ангилан зохицуулах энэхүү механизм нь бүх төрлийн AI системд нэг ижил шаардлага тавих бус, харин түүний бодит эрсдэлийн түвшинд тохируулан ялгавартай зохицуулалт хийх боломжийг бүрдүүлж байгаагаараа онцлог юм. Энэ нь нэг талаас инновацыг дэмжих, нөгөө талаас хүний эрхийг хамгаалах тэнцвэрийг хангах эрх зүйн чухал шийдэл болж байна.

Гурав. AI-ийн хэрэглээнээс үүдэх хүний эрхийн эрсдэл

AI-ийн хэрэглээ нь мэдээлэл боловсруулах, шийдвэр гаргах үйл явцыг автоматжуулж, үр ашигтай болгох давуу талтай боловч түүний практик хэрэглээ нь хүний үндсэн эрх, эрх чөлөөнд бодит эрсдэл учруулж байна. Ялангуяа алгоритмд суурилсан шийдвэр гаргалт нь хүний оролцоог хязгаарлах, шийдвэр гаргах үйл явцын ил тод байдлыг бууруулах, улмаар эрх зүйн хамгаалалтын уламжлалт баталгааг сулруулах нөхцөлийг бүрдүүлж байна.

1. Ялгаварлан гадуурхах эрсдэл (Diskriminierung)

AI системүүд нь өгөгдөлд агуулагдаж буй тэгш бус байдал, хэвшмэл ойлголтыг давтан бий болгох, улмаар тодорхой бүлгийг ялгаварлан гадуурхах шийдвэр гаргах эрсдэлтэй байдаг[12].

Энэ эрсдэлийг бодитоор харуулсан жишээ нь Amazon компанийн ажилд авах AI системтэй холбоотой тохиолдол юм. Тус компани 2014–2018 оны хооронд ажилд авах процессыг автоматжуулах зорилгоор AI-д суурилсан систем ашигласан бөгөөд уг систем нь өмнөх жилүүдийн ажилд авсан ажилтнуудын өгөгдөлд тулгуурлан нэр дэвшигчдийг үнэлдэг байсан байна.

Гэвч уг өгөгдөл нь ихэвчлэн эрэгтэй ажилтнууд давамгайлсан бүтэцтэй байсан тул AI систем эмэгтэй нэр дэвшигчдийг системтэйгээр доогуур үнэлж, зарим тохиолдолд шууд хасах хандлагатай болсон нь тогтоогдсон. Үүний улмаас тус компани уг AI

системийг хэрэглэхээс татгалзсан байна[13].

Энэхүү жишээ нь AI систем нь “саармаг” мэт харагдах боловч бодит байдал дээр өмнөх тэгш бус байдлыг давтан бий болгож, ялгаварлан гадуурхах шийдвэр гаргах эрсдэлтэйг тод харуулж байна.

Иймд AI ашиглан ажилд авах, зээл олгох, эрүүгийн хэрэгт эрсдэлийн үнэлгээ хийх зэрэг салбарт гаргасан шийдвэр нь тодорхой этгээдийг үндэслэлгүйгээр сөргөөр үнэлэх, боломжийг хязгаарлах нөхцөлийг бүрдүүлж болзошгүй бөгөөд энэ нь эрх тэгш байдлын зарчимыг зөрчих эрсдэлийг үүсгэнэ.

2. Хувийн нууц, мэдээлэл хамгаалалтын эрсдэл (Privatsphäre, Datenschutz)

AI системүүд нь их хэмжээний өгөгдөлд тулгуурлан ажилладаг тул хувь хүний мэдээллийг өргөн хүрээнд цуглуулах, боловсруулах шаардлага үүсдэгийн улмаас хувь хүний нууцад халдах, зөвшөөрөлгүйгээр мэдээлэл ашиглах, эсхүл зорилгын хязгаарлалтыг зөрчих эрсдэл бий болдог[14]. Ялангуяа профайлчлал болон автомат шийдвэр гаргалт нь тухайн этгээдийн эрх, хууль ёсны ашиг сонирхолд шууд нөлөөлөх шинжтэй бөгөөд хүний эрхийн хамгаалалтын хүрээнд онцгой анхаарах шаардлагатай асуудалд тооцогдож байна.

Энэхүү эрсдэлийг бодитоор харуулсан жишээ нь Clearview AI компанийн үйл ажиллагаа бөгөөд тус компани интернэт орчин, олон нийтийн сүлжээнээс иргэдийн зургийг программын аргаар системтэйгээр цуглуулж, улмаар тэрбум гаруй зургийг биометрийн өгөгдөл болгон боловсруулж, нүүр таних хиймэл оюуны системд ашиглах дэлхий даяарх зураг мэдээллийн сан бүрдүүлэн, хувь хүнийг зөвшөөрөлгүйгээр таних, зан үйлийг хянах боломж бүхий системийг бий болгосон байна[15].

Уг систем нь хэрэглэгч (ялангуяа хууль сахиулах байгууллага) тодорхой нэг хүний зургийг оруулахад тухайн хүний интернэт дэх бусад зураг, холбогдох мэдээллийг илрүүлэх боломж олгодог бөгөөд энэ нь бодит байдал дээр иргэдийг зөвшөөрөлгүйгээр таних, хянах нөхцөлийг бүрдүүлсэн.

Үүнтэй холбогдуулан Европын холбооны мэдээлэл хамгаалах байгууллагууд уг үйл ажиллагааг хувь хүний мэдээлэл хамгаалах эрхийг зөрчсөн гэж дүгнэсэн бөгөөд тухайлбал Италийн мэдээлэл хамгаалах байгууллага болох Garante per la protezione dei dati personali 2022 онд Clearview AI компанид 20 сая еврогийн торгууль ногдуулж, Италийн иргэдийн өгөгдлийг цуглуулах, боловсруулахыг хориглон, өмнө цуглуулсан мэдээллийг устгах шийдвэр гаргасан байна[16].

Энэхүү тохиолдол нь AI систем нь хувь хүний мэдээллийг зөвшөөрөлгүйгээр, ил тод бус аргаар цуглуулах, боловсруулах замаар хувийн нууцад ноцтой халдах эрсдэлтэйг тод харуулж байна. Мөн уг үйл ажиллагаа нь хувь хүний мэдээллийг тодорхой, хууль ёсны зорилгоор хязгаарлагдмал байдлаар ашиглах ёстой гэсэн зарчим, түүнчлэн мэдээлэл боловсруулах ажиллагаа нь тухайн этгээдийн мэдлэг, зөвшөөрөлд суурилсан байх шаардлагыг зөрчиж байгаагаараа онцлог юм.

Иймд AI системийн хүрээнд хувь хүний мэдээллийг боловсруулах ажиллагаанд ил тод байдал, зөвшөөрөл, зорилгын хязгаарлалт, мөн хараат бус хяналтын механизмыг

хангах нь хүний эрхийн хамгаалалтын зайлшгүй баталгаа болох шаардлагатай байна.

3. Шударга шүүхээр шүүлгэх эрхийн баталгаанд үзүүлэх эрсдэл[17] (Recht auf ein faires Verfahren)

Дээр дурдсан эрсдэлүүд нь зөвхөн мэдээлэл боловсруулах шатанд бус, шүүхийн ажиллагаанд хүртэл нөлөөлөх боломжтой.

AI системийг хууль сахиулах болон шүүхийн ажиллагаанд ашиглах нь нотлох баримтын үнэлгээ, эрх зүйн дүгнэлт боловсруулахад дэмжлэг үзүүлэх боломжтой боловч практикт AI-аас гаргасан мэдээлэл, хариултад хэт итгэх хандлага нь шударга шүүхээр шүүлгэх эрхийн баталгаанд сөргөөр нөлөөлөх эрсдэлийг бий болгож байна.

Тухайлбал, Mata v. Avianca, Inc[18]. хэргийн хүрээнд АНУ-ын Нью-Йорк мужийн Өмнөд дүүргийн Холбооны шүүхэд (United States District Court for the Southern District of New York) хянан хэлэлцэгдсэн иргэний хэрэгт өмгөөлөгч AI систем (ChatGPT)-ээс авсан шүүхийн шийдвэрүүдийг бодит эх сурвалж гэж үзэн шүүхэд бичгээр гаргасан байна.

Гэвч шүүхээс уг иш татсан шийдвэрүүдийг шалгахад тэдгээр нь бодитоор оршдоггүй, өөрөөр хэлбэл AI-аас “зохиомлоор үүсгэсэн” хуурамч шийдвэрүүд болох нь тогтоогдсон. Үүний улмаас шүүхээс тухайн өмгөөлөгч болон түүний хуулийн фирмд мөнгөн торгууль ногдуулж, мэргэжлийн хариуцлага хүлээлгэсэн байна.

Энэхүү тохиолдол нь AI-аас гаргасан мэдээллийг шалгалгүйгээр ашиглах нь шүүхийн ажиллагаанд буруу мэдээлэл нэвтрүүлэх, улмаар нотлох баримтын үнэн зөв байдал, шийдвэрийн хууль ёсны үндэслэлд ноцтой сөрөг нөлөө үзүүлэх эрсдэлтэйг тод харуулж байна. Ийм нөхцөлд тухайн этгээд шийдвэрийн үндэслэлийг ойлгох, эсэргүүцэх, өөрийгөө хамгаалуулах эрх нь хязгаарлагдах эрсдэл үүсэж, шударга шүүхээр шүүлгэх эрхийн үндсэн баталгаанд харшилж болзошгүй нь шударга шүүхээр шүүлгэх эрхийн үндсэн баталгаанд харшлах эрсдэлийг бий болгож байна.

4. Хяналтгүй автомат шийдвэр гаргалтын эрсдэл (automatisierte Entscheidungsfindung)

AI-д суурилсан системүүд нь хүний шууд оролцоогүйгээр шийдвэр гаргах боломжтой бөгөөд энэ нь хариуцлагын тодорхой бус байдал, хяналтын сулралд хүргэх бодит эрсдэлийг бий болгож байна. Ялангуяа автомат шийдвэр гаргалт нь шийдвэр гаргах үйл явцад оролцсон субъект, хариуцлага хүлээх этгээдийг тодорхойлох, алдаа гарсан тохиолдолд түүнийг засах, хяналт тавих механизмыг хэрэгжүүлэхэд хүндрэл учруулж, улмаар эрх зүйн баталгааг сулруулах нөхцөлийг бүрдүүлнэ.

Энэхүү эрсдэлийг Европт бодитоор харуулсан жишээ нь Нидерланд улсад хэрэгжүүлж байсан SyRI (System Risk Indication) системтэй холбоотой хэрэг юм. Уг систем нь төрийн байгууллагуудаас иргэдийн олон төрлийн мэдээллийг нэгтгэн автомат аргаар боловсруулж, халамжийн зөрчлийн эрсдэлийг тодорхойлох зорилготой байсан боловч алгоритмын ажиллагаа ил тод бус, иргэдийн талаар ямар шалгуураар дүгнэлт гаргаж байгааг тайлбарлах боломжгүй байсан нь шүүмжлэлд өртсөн байна[19].

Улмаар 2020 онд SyRI Urteil шийдвэрээр Нидерландын шүүх уг системийг хүний эрх, тухайлбал хувийн нууц болон шударга ажиллагааны зарчимд харшилсан гэж үзэн

хэрэглэхийг хориглосон байна[20].

Энэхүү тохиолдол нь автомат шийдвэр гаргалт нь хүний хяналтгүй, ил тод бус байдлаар хэрэгжсэн нөхцөлд зөвхөн алдаа гаргах эрсдэлтэй бус, харин тухайн шийдвэрийн үндэслэл, хяналт, хариуцлагын механизм тодорхойгүй болж, хүний эрхийн хамгаалалтын баталгааг сулруулах аюултайг харуулж байна.

Хүний хяналт хангалтгүй нөхцөлд буруу, эсхүл шударга бус шийдвэр гарсан ч түүнийг цаг тухайд нь илрүүлэх, засах, хариуцлага тооцох боломж хязгаарлагдаж, улмаар тухайн этгээдийн эрхийг хамгаалах баталгаа алдагдах эрсдэл үүсдэг.

Иймд AI-д суурилсан автомат шийдвэр гаргалтыг хэрэглэхдээ хүний хяналтыг зайлшгүй хангах, шийдвэр гаргалтын ил тод байдлыг нэмэгдүүлэх, хариуцлагын тодорхой механизмыг тогтоох нь хүний эрхийг хамгаалах, төрийн байгууллагын үйл ажиллагааны хууль ёсны байдлыг хангах үндсэн нөхцөл болох шаардлагатай байна.

Ийнхүү AI-ийн хэрэглээнээс үүдэх дээрх эрсдэлүүд нь эрх тэгш байдал, хувийн нууц, шударга шүүхээр шүүлгэх эрх, ил тод байдал, хариуцлагын зарчмуудтай шууд холбогдож байна.

Энэ нь AI-ийг хүний оролцоог орлох хэрэгсэл бус, харин хүний шийдвэр гаргалтыг дэмжих, заавал хүний хяналттайгаар ашиглах шаардлагатайг харуулж байна.

Дөрөв. Хувь хүний мэдээлэл хамгаалах эрхийн баталгаа ба AI (DSGVO)[21]

AI-ийн хэрэглээ өргөжин тэлж, хувь хүний мэдээллийг өргөн хүрээнд боловсруулах болсон нөхцөлд мэдээлэл хамгаалалтын эрх зүйн баталгааг тодорхой болгох шаардлага үүсээд байна[1]. Энэ хүрээнд “Европын холбооны ерөнхий мэдээлэл хамгаалах журам” (DSGVO) нь AI-ийн хүрээнд хувь хүний мэдээллийг хамгаалах үндсэн эрх зүйн суурийг бүрдүүлж байна.

DSGVO нь хувь хүний мэдээллийг боловсруулахад баримтлах үндсэн зарчмуудыг тогтоосон бөгөөд эдгээр нь AI системийн үйл ажиллагаанд шууд хамааралтай. Үүнд:

1. Зөвшөөрөл

DSGVO-ийн дагуу хувь хүний мэдээллийг боловсруулах үндсэн хууль ёсны суурь нь тухайн этгээдийн урьдчилсан, тодорхой, мэдээлэлд суурилсан, сайн дурын зөвшөөрөл байх ёстой. Өөрөөр хэлбэл, мэдээлэл боловсруулах үйл ажиллагаа нь тухайн этгээд өөрийн мэдээлэл ямар зорилгоор, ямар хүрээнд ашиглагдахыг бүрэн ойлгосон, чөлөөтэй илэрхийлсэн зөвшөөрөлд үндэслэх шаардлагатай.

Гэвч AI системүүдийн хувьд, ялангуяа их хэмжээний өгөгдөлд тулгуурласан орчинд мэдээллийг гуравдагч эх үүсвэрээс цуглуулах, дахин ашиглах, нэгтгэх зэрэг практик өргөн тархсан нь зөвшөөрлийн зарчмыг бодитоор хэрэгжүүлэхэд хүндрэл учруулж байна. Тухайлбал, Clearview AI компанийн жишээнээс харахад интернэт орчин дахь иргэдийн зургийг урьдчилсан зөвшөөрөл авалгүйгээр цуглуулж, биометрийн өгөгдөл болгон боловсруулсан нь DSGVO-ийн шаардлагад нийцэхгүй болохыг Европын мэдээлэл хамгаалах байгууллагууд тогтоосон байна[22].

Энэхүү нөхцөл байдал нь AI системийн хүрээнд зөвшөөрөл нь формаль буюу цаасан дээрх шаардлага төдий бус, харин бодитой, хэрэгжих боломжтой, хяналттай байх ёстойг харуулж байна. Иймд AI-д суурилсан мэдээлэл боловсруулах ажиллагаанд зөвшөөрлийг ил тод, ойлгомжтой хэлбэрээр авах, мөн тухайн этгээд зөвшөөрлөө цуцлах боломжийг бодитоор хангах нь эрх зүйн зайлшгүй шаардлага болно.

2.Зорилгын хязгаарлалт[23]

DSGVO-ийн дагуу хувь хүний мэдээллийг зөвхөн тодорхой, хууль ёсны, урьдчилан тодорхойлогдсон зорилгын хүрээнд цуглуулж, тухайн зорилготой нийцэх байдлаар боловсруулах ёстой бөгөөд үүнийг зорилгын хязгаарлалтын зарчим гэж ойлгоно. Өөрөөр хэлбэл, мэдээллийг нэг зорилгоор цуглуулсан бол тухайн зорилгоос өөр, нийцэхгүй зорилгоор ашиглахыг хориглодог.

Гэвч AI системүүдийн онцлог нь олон эх үүсвэрээс өгөгдөл цуглуулах, тэдгээрийг нэгтгэн дахин боловсруулах, шинэ зорилгоор ашиглах боломжийг бүрдүүлдэгт оршиж байна. Энэ нь практикт мэдээллийг анх цуглуулсан зорилгоос өөр зорилгоор ашиглах, өөрөөр хэлбэл дээрх зарчмыг зөрчих эрсдэлийг бий болгож байна.

Тухайлбал, Clearview AI компанийн жишээнээс харахад иргэдийн сошиал сүлжээнд байршуулсан зургууд нь анх харилцаа, мэдээлэл солилцох зорилготой байсан боловч уг мэдээллийг нүүр таних хиймэл оюуны системд ашиглах зорилгоор дахин боловсруулсан нь анхны зорилготой нийцэхгүй үйлдэлд тооцогдсон байна[24].

Энэхүү нөхцөл байдал нь AI системийн хүрээнд мэдээлэл боловсруулах ажиллагаа нь динамик, олон зорилготой болж өөрчлөгдөж байгаа үед “Зорилгын хязгаарлалт” зарчмыг хэрэгжүүлэхэд бодит хүндрэл үүсэж байгааг харуулж байна. Ялангуяа их өгөгдөлд суурилсан орчинд мэдээллийг “ирээдүйн тодорхойгүй зорилгоор” цуглуулах хандлага нь уг зарчмын суурь утгатай зөрчилдөж байна.

Иймд AI-д суурилсан мэдээлэл боловсруулах ажиллагаанд зорилгын хязгаарлалтыг хангахын тулд мэдээлэл цуглуулах үе шатанд зорилгыг тодорхой, нарийвчилсан байдлаар тогтоох, мөн тухайн зорилгоос хэтэрсэн ашиглалтыг хязгаарлах эрх зүйн болон техникийн механизм хэрэгжүүлэх шаардлагатай байна.

3. Автомат шийдвэр гаргалтад хязгаарлалт[25]

DSGVO-ийн 22 дугаар зүйлд (Art. 22 DSGVO) хувь хүнийг зөвхөн автоматжуулсан шийдвэр гаргалтад үндэслэн, түүнд хууль эрх зүйн үр дагавар үүсгэх, эсхүл түүнд ижил төстэй ноцтой нөлөө үзүүлэх шийдвэрт өртөхгүй байх эрхийг баталгаажуулсан байдаг.

Өөрөөр хэлбэл, тухайн этгээдийн эрх, хууль ёсны ашиг сонирхолд ноцтой нөлөөлөх шийдвэрийг бүрэн автомат системд даатгахыг ерөнхийдөө хориглож, зайлшгүй тохиолдолд хүний оролцоо, тайлбар авах боломж, мөн шийдвэрийг эсэргүүцэх эрхийг хангах шаардлагыг тавьдаг.

Гэвч AI системүүдийн хөгжлийн өнөөгийн түвшинд, ялангуяа алгоритмын шийдвэр гаргалт нь өндөр хурд, үр ашигтай байдлаар хэрэгжиж байгаа нөхцөлд автомат

шийдвэр гаргалтыг бодитоор хязгаарлах, хүний оролцоог үр дүнтэй хангах асуудал практикт хүндрэлтэй хэвээр байна.

Тухайлбал, Нидерланд улсад хэрэгжүүлж байсан SyRI (System Risk Indication) систем нь төрийн байгууллагууд иргэдийн мэдээллийг нэгтгэн автомат аргаар боловсруулж, халамжийн зөрчил илрүүлэх зорилгоор эрсдэлийн үнэлгээ хийдэг байсан боловч уг системийн алгоритмын ажиллагаа ил тод бус, тухайн иргэн ямар шалгуураар “эрсдэлтэй” гэж үнэлэгдэж байгааг ойлгох боломжгүй байсан нь шүүмжлэлд өртсөн байна. Улмаар 2020 онд SyRI Urteil шийдвэрээр уг системийг хүний эрх, тухайлбал хувийн нууц болон шударга ажиллагааны зарчимд харшилсан гэж үзэн хэрэглэхийг хориглосон[26].

Энэхүү тохиолдол нь автомат шийдвэр гаргалт нь хүний оролцоогүй, ил тод бус байдлаар хэрэгжсэн нөхцөлд зөвхөн алдаа гаргах эрсдэлтэй бус, харин тухайн шийдвэрийн үндэслэл, хяналт, хариуцлагын механизм тодорхойгүй болж, хувь хүний эрхийг хамгаалах боломжийг бодитоор хязгаарлаж байгааг харуулж байна.

Иймд Art. 22 DSGVO-д тусгагдсан зохицуулалт нь AI-ийн хэрэглээний хүрээнд зөвхөн формаль хязгаарлалт бус, харин хүний оролцоог бодитоор хангах, шийдвэр гаргалтыг тайлбарлах боломжийг нэмэгдүүлэх, мөн тухайн этгээдийн эрхийг хамгаалах бодит механизм болгон хэрэгжүүлэх шаардлагатай байна.

4. Ил тод байдал ба тайлбарлах боломж[27]

DSGVO-ийн дагуу хувь хүний мэдээлэл боловсруулах ажиллагаа нь ил тод байх, тухайн этгээд өөрийн мэдээлэл хэрхэн, ямар зорилгоор, ямар үндэслэлээр ашиглагдаж байгааг мэдэх эрхтэй байх ёстой. Энэ нь мэдээлэл боловсруулах ажиллагаа нь зөвхөн хууль ёсны үндэслэлтэй байх төдийгүй, тухайн этгээдэд ойлгомжтой, хяналттай байх шаардлагыг илэрхийлдэг.

AI системийн хувьд энэхүү шаардлага нь шийдвэр гаргалтын үндэслэлийг тайлбарлах боломжтой салшгүй холбоотой. Өөрөөр хэлбэл, AI-аас гаргасан шийдвэр, дүгнэлт нь ямар өгөгдөл, ямар логикт тулгуурласан болохыг тодорхой хэмжээнд тайлбарлах боломжтой байх шаардлагатай.

Гэвч практикт олон AI системүүд “хар хайрцаг”-ийн шинжтэй буюу дотоод алгоритмын ажиллагаа, шийдвэр гаргах логик нь хэрэглэгч болон хяналт хэрэгжүүлэгч байгууллагад ойлгомжгүй, тайлбарлах боломж хязгаарлагдмал байдаг[28]. Энэ нь шийдвэрийн үндэслэлийг шалгах, үнэлэх, эсэргүүцэх боломжийг хязгаарлаж, улмаар хүний эрхийн хамгаалалтын баталгаанд сөргөөр нөлөөлөх эрсдэлийг бий болгодог.

Тухайлбал, автомат шийдвэр гаргалтад тулгуурласан эрсдэлийн үнэлгээ, кредит оноо тогтоох, эсхүл төрийн халамжийн шалгуур тогтоох системүүдийн хувьд тухайн этгээд өөрийг нь ямар шалгуураар үнэлж, ямар үндэслэлээр шийдвэр гаргасныг ойлгох боломжгүй байх нь түгээмэл ажиглагдаж байна. Энэ нь зөвхөн ил тод байдлын зарчмыг зөрчихөөс гадна тухайн этгээдийн эрхээ хамгаалах, шийдвэрийг эсэргүүцэх боломжийг бодитоор хязгаарлаж байна.

Иймд AI системийн хүрээнд ил тод байдлыг хангахын тулд мэдээлэл боловсруулах

ажиллагааны талаар тодорхой мэдээлэл өгөх, мөн тайлбарлах боломжыг нэмэгдүүлэх техникийн болон эрх зүйн механизм хэрэгжүүлэх шаардлагатай байна. Энэ нь зөвхөн мэдээлэл хамгаалалтын шаардлага төдийгүй, хүний эрхийн хамгаалалтын үндсэн баталгаа болох ач холбогдолтой.

Ийнхүү “Европын холбооны ерөнхий мэдээлэл хамгаалах журам”-д тусгагдсан зөвшөөрөл, зорилгын хязгаарлалт, автомат шийдвэр гаргалтын хязгаарлалт (Art. 22 DSGVO), ил тод байдал болон тайлбарлах боломж зэрэг зарчмууд нь AI-ийн хэрэглээтэй холбоотой хувь хүний мэдээлэл хамгаалах, хүний эрхийг баталгаажуулах үндсэн эрх зүйн суурийг бүрдүүлж байна.

Гэсэн хэдий ч AI системийн технологийн онцлог, тухайлбал их хэмжээний өгөгдөлд тулгуурлах байдал, алгоритмын төвөгшил, “хар хайрцаг”-ийн шинж, мөн автомат шийдвэр гаргалтын өндөр түвшин нь дээрх зарчмуудыг практикт бүрэн хэрэгжүүлэхэд тодорхой хүндрэл, сорилтыг бий болгож байна.

Иймд DSGVO-д тусгагдсан зохицуулалтыг AI-ийн хэрэглээний онцлогт нийцүүлэн зөвхөн формаль шаардлага төдий бус, харин бодит хэрэгжилтийг хангахуйц, үр дүнтэй хяналт, хэрэгжилтийн механизмтай уялдуулан хэрэгжүүлэх шаардлагатай байна. Тухайлбал, хүний хяналтыг бодитоор хангах, мэдээлэл боловсруулах ажиллагааны ил тод байдлыг нэмэгдүүлэх, мөн AI системийн шийдвэр гаргалтыг тайлбарлах боломжийг сайжруулах нь хүний эрхийн хамгаалалтын зайлшгүй нөхцөл болох юм.

Энэ нь эцсийн дүндээ AI-ийн зохицуулалтыг зөвхөн мэдээлэл хамгаалалтын хүрээнд бус, харин хүний эрхийн иж бүрэн баталгааг хангах эрх зүйн механизмын салшгүй хэсэг болгон авч үзэх шаардлагатайг илтгэж байна.

Тав. Нэгдсэн дүгнэлт, санал

Хиймэл оюун ухаан (AI)-ийн хэрэглээ нь мэдээлэл боловсруулах, шийдвэр гаргалтыг хурдасгах, үр ашгийг нэмэгдүүлэх зэрэг давуу талтай боловч түүний практик хэрэглээ нь хүний эрхийн үндсэн зарчмуудад бодит эрсдэл учруулж байгааг дээрх судалгаа, жишээнүүд тодорхой харуулж байна.

Ялангуяа AI системүүд нь:

Мэдээллийн өгөгдөлд агуулагдсан тэгш бус байдлыг давтан бий болгож ялгаварлан гадуурхах эрсдэл үүсгэх,

Хувь хүний мэдээллийг зөвшөөрөлгүйгээр цуглуулах, зорилгын хязгаарлалтыг зөрчих,

Шүүхийн ажиллагаанд буруу, баталгаагүй мэдээлэл нэвтрүүлэх замаар шударга шүүхээр шүүлгэх эрхэд сөргөөр нөлөөлөх,

Хүний оролцоогүй автомат шийдвэр гаргалтын улмаас хяналт, хариуцлагын механизм сулрах

зэрэг эрсдэлийг бий болгож байна.

Үүнтэй холбоотойгоор Европын холбооноос “эрсдэлд суурилсан хандлага”-д тулгуурласан AI-ийн зохицуулалт, мөн (DSGVO) “Европын холбооны ерөнхий мэдээлэл хамгаалах журам”-аар дамжуулан хувь хүний мэдээлэл хамгаалах эрх зүйн баталгааг бий болгосон нь чухал ач холбогдолтой боловч AI-ийн технологийн онцлог, тухайлбал алгоритмын төвөгшил, их өгөгдөлд суурилсан боловсруулалт зэрэг нь эдгээр зохицуулалтыг практикт бүрэн хэрэгжүүлэхэд хүндрэл учруулж байгааг харуулж байна.

Иймд AI-ийн хэрэглээг хүний эрхийн баталгаатай уялдуулан зохицуулах хүрээнд дараах санал, зөвлөмжийг дэвшүүлж байна:

1. AI системийн эрсдэлийг үнэлэх, ангилах механизмыг үндэсний эрх зүйн зохицуулалтад нэвтрүүлж, ялангуяа өндөр эрсдэлтэй системд хатуу хяналт тогтоох;

2. AI-д суурилсан шийдвэр гаргалтад хүний хяналтыг зайлшгүй хангах, автомат шийдвэрийг шалгах, засах, эсэргүүцэх бодит механизмыг бүрдүүлэх;

3. Хувь хүний мэдээлэл боловсруулах ажиллагаанд зөвшөөрөл, зорилгын хязгаарлалт, ил тод байдлын зарчмыг бодитоор хэрэгжүүлэх, формаль хэрэгжилтээс зайлсхийх;

4. AI системийн шийдвэр гаргалтын ил тод байдал, тайлбарлах боломжийг нэмэгдүүлэх, “хар хайрцаг”-ийн шинжийг бууруулах техникийн болон эрх зүйн шаардлагыг хөгжүүлэх;

5. AI-ийн хэрэглээтэй холбоотой хүний эрхийн эрсдэлийн талаар мэргэжилтнүүдийн мэдлэг, ур чадварыг дээшлүүлэх сургалт, арга хэмжээг тогтмол зохион байгуулах.

Эцэст нь, AI нь хүний шийдвэр гаргалтыг орлох бус, харин дэмжих хэрэгсэл байх ёстой бөгөөд түүнийг хүний эрхийн баталгаатай уялдуулан, хяналттай, ил тод байдлаар ашиглах нь орчин үеийн эрх зүйн хөгжлийн зайлшгүй чиг хандлага болохыг энэхүү судалгаа харуулж байна.

[1] European Commission, Artificial Intelligence Act – Overview. https://artificial-intelligence-act.eu/

[2] BfDI, Künstliche Intelligenz und Datenschutz. https://www.bfdi.bund.de/DE/Buerger/Inhalte/Technik/Kuenstliche-Intelligenz.html

[3] Max Planck Institute for Innovation and Competition, Artificial Intelligence and Law. https://www.ip.mpg.de/en/research/research-news/artificial-intelligence-and-law.html

[4] European Commission, Proposal for a Regulation laying down harmonised rules on artificial intelligence, 2021. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

[5] European Commission, Risk-based approach under the AI Act. https://artificial-intelligence-act.eu/the-act/

[6] Europäische Kommission, Gesetz über künstliche Intelligenz (AI Act) – Überblick.

https://artificial-intelligence-act.eu/de/

[7] Europäische Kommission, Risikobasierter Ansatz im AI Act. https://artificial-intelligence-act.eu/the-act/

[8] Europäische Kommission, Verbotene Praktiken im Bereich der künstlichen Intelligenz. https://artificial-intelligence-act.eu/prohibited-ai-practices/

[9] Europäische Kommission, Hochrisiko-KI-Systeme. https://artificial-intelligence-act.eu/high-level-summary/

[10] Europäische Kommission, Transparenzanforderungen für KI-Systeme. https://artificial-intelligence-act.eu/transparency/

[11] Europäische Kommission, KI-Systeme mit minimalem Risiko.https://artificial-intelligence-act.eu/

[12] Max Planck Institute for Innovation and Competition, Diskriminierungsrisiken durch Künstliche Intelligenz. https://www.ip.mpg.de

[13] Reuters, Amazon schafft KI-Rekrutierungssystem ab, das Frauen benachteiligte, 2018. https://www.reuters.com/article/us-amazon-jobs-automation-insight-idUSKCN1MK08G

[14] BfDI, Künstliche Intelligenz und Datenschutz. https://www.bfdi.bund.de/DE/Buerger/Inhalte/Technik/Kuenstliche-Intelligenz.html

[15] Heise Online, Gesichtserkennung von Clearview: Milliarden Fotos aus dem Internet gesammelt, 2025. https://www.heise.de/news/UK-Gesichtserkennung-von-Clearview-unterliegt-europaeischen-Datenschutzregeln-10752341.html

[16] Datenschutzkonferenz, Einsatz von Gesichtserkennungssystemen durch Clearview AI rechtswidrig, 2022. https://www.datenschutzkonferenz-online.de

[17] AI-ийг хуульчид ашиглахад тавигдах шаардлага, эрсдэлийг тайлбарласан. Bundesrechtsanwaltskammer, Hinweise zum Einsatz von Künstlicher Intelligenz in der anwaltlichen Praxis, 2023. https://www.brak.de

[18] United States District Court, Southern District of New York, Mata v. Avianca, Inc., Nr. 22-cv-1461 (PKC), Beschluss vom 22. Juni 2023. https://storage.courtlistener.com/recap/gov.uscourts.nysd.575368/gov.uscourts.nysd.575368.54.0.pdf

[19] European Digital Rights, Das SyRI-System und seine Auswirkungen auf die Grundrechte, 2020.https://edri.org

[20] Bezirksgericht Den Haag, Urteil zum SyRI-System, 05.02.2020.

[21] Datenschutz-Grundverordnung, Art. 6 Abs. 1 lit. a DSGVO.

[22] Garante per la protezione dei dati personali, Clearview AI: Verstoß gegen die DSGVO wegen fehlender Einwilligung, 2022.

[23] Datenschutz-Grundverordnung, Art. 5 Abs. 1 lit. b DSGVO.

[24] Garante per la protezione dei dati personali, Clearview AI: Verstoß gegen den Grundsatz der Zweckbindung, 2022.

[25] Datenschutz-Grundverordnung, Art. 22 DSGVO.

[26] Bezirksgericht Den Haag, Urteil zum SyRI-System, 05.02.2020.

[27] Datenschutz-Grundverordnung, Art. 12–14 DSGVO (Transparenzpflichten).

[28] Max Planck Institute for Innovation and Competition, Transparenz und Erklärbarkeit von KI-Systemen, 2022.

Эх сурвалж: https://prosecutor.mn/post/709