

5. modulis

Uz vērtībām balstīta
projektēšana mākslīgā
intelekta inženierijā



Co-funded by
the European Union



Ethical Engineer

www.ethicalengineer.eu

SATURS

- 01 Iepazīstieties ar ieinteresētajām pusēm un vērtībām
- 02 Vērtību pārvēršana dizainā
- 03 Salīdziniet vērtības dažādos gadījumos
- 04 Aizstāvēt uz vērtībām balstītu lēmumu



Co-funded by
the European Union

Šis projekts ir finansēts ar Eiropas Komisijas atbalstu. Par šo publikāciju (paziņojumu) atbild vienīgi autors, un Komisija neuzņemas nekādu atbildību par tajā ietvertās informācijas izmantošanu. Saskaņā ar jauno Vispārīgās datu aizsardzības regulas (GDPR) regulējumu lūdz ņemiet vērā, ka partnerība apstrādās jūsu personas datus vienīgi projekta interesēs un nolūkos, neierobežojot jūsu tiesības.



Īss pārskats par kursu „Ētiskais inženieris”

Katras moduļa galvenais saturs un mācību rezultāti

1. modulis	02. modulis	03. modulis	04. modulis	05. modulis	06. modulis
Ētika un pārredzamība	Priekšstati, taisnīgums un iekļaušana	Privātums un datu aizsardzība	Atbildība un drošība	Uz vērtībām balstīts dizains	Ilgspēja un ietekme
Ētikas principi un ES vērtības Pārredzamība un izskaidrojamība AI un robotikas gadījumu analīze	Priekšstati visā AI dzīves ciklā Taisnīguma mērķi un pierādījumi Iekļaujošas inženierijas gadījumi	Privātums jau projektēšanas posmā un Vispārīgā datu aizsardzības regula (GDPR) Datu pārvaldība un PET Novērtēšana un apstrīdējamība	Lomas un atbildība Kļūdas, noturība un nodošana Reģistrēšana un mācīšanās no incidentiem	Ētiskās dizaina prasības Datu, modeļu un saskarnes izvēle Kompromisi un ātrās pārbaudes	Energija, ūdens un aparatūra Ietekme uz sabiedrību un dzīves ciklu Rādītāji, pārvaldība un pārveidošana

Moduļa ievads

Šis modulis pārvērs ieinteresēto pušu vērtības prasībās, projektēšanas alternatīvās, testēšanas rezultātos un pamatotā inženierijas lēmumā.

Mērķis

Sniegt inženierzinātņu studentiem iespēju izpētīt vērtības, pārvērst tās projektēšanas kritērijos un pamatot savu izvēli, balstoties uz ieinteresēto pušu viedokli un tehniskajiem pierādījumiem.

Mērķauditorija

Augstskolu studenti inženierzinātnēs, datorzinātnēs, produktu dizainā, robotikā un saistītās disciplīnās.

Pamatojas uz

1.–4. modulis: ētika, pārredzamība, taisnīgums, privātums, atbildība, drošība un izsekojami inženierijas lēmumi.

[Avoti]

- „Ethical Engineer” apstiprināja KA220-HED pieteikumu, WP3 „Mācīšana un mācīšanās”.
- https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf
- <https://standards.ieee.org/standard/7000-2021.html>

Moduļu orientācija

Šis modulis pārvērs ieinteresēto pušu vērtības prasībās, projektēšanas alternatīvās, testēšanas rezultātos un pamatotā inženierijas lēmumā.

Mācību secība

Atklāt ieinteresēto pušu vērtības, pārvērst tās projektā, salīdzināt gadījumus un pamatot uz vērtībām balstītu lēmumu.

Mācību rezultāti

Ieinteresēto pušu karte, vērtību hierarhija, prasības, variantu salīdzinājums un uz vērtībām balstīts projektēšanas protokols.

[Avoti]

- „Ethical Engineer” apstiprināja KA220-HED pieteikumu, WP3 „Mācīšana un mācīšanās”.
- https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf
- <https://standards.ieee.org/standard/7000-2021.html>

Mācību rezultāts

Līdz moduļa
beigām studenti
spēs:

1. Identificēt tiešās un netiešās ieinteresētās puses, varas atšķirības un attiecīgās vērtības.
2. veikt konceptuālus, empīriskus un tehniskus pētījumus par vērtībām atbilstošu dizainu.
3. pārvērst vērtības prasībās, pieņemšanas kritērijos un izsekojamā dizaina pierādījumos.
4. salīdzināt datu, modeļu, saskarnes un sistēmas variantus, izmantojot sliekšņus, pierādījumus un nenoteiktību.
5. izveidot uz vērtībām balstītu projektēšanas dokumentāciju, iekļaujot drošības pasākumus, atlikušās pretrunas un pārskatīšanas nosacījumus.

[Avoti]

- „Ethical Engineer” apstiprināja KA220-HED pieteikumu, WP3 „Mācīšana un mācīšanās”.
- https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf
- <https://standards.ieee.org/standard/7000-2021.html>
- <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

Vērtības jau ir klātesošas inženierijas izvēlēs

- Mērķi nosaka panākumus. Datu kopas nosaka, kuras pieredze ir redzama. Sliekšņi sadala risku. Saskarnes veido uzmanību un kontroli.
- AI inženierija ir sociotehniska: uzvedība veidojas no cilvēkiem, datiem, modeļiem, saskarnēm, organizācijām un darbības konteksta.
- Uz vērtībām balstīts dizains padara šīs izvēles skaidras, uz pierādījumiem balstītas un pārskatāmas. Inženierijas uzdevums ir parādīt, kā vērtība ietekmēja lēmumu un kā tiks pārbaudīta prasība.

[Avoti]

- Oriģinālā 5. moduļa prezentācija, ko nodrošinājis projekta „Ethical Engineer” konsorcijs.
- https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf
- <https://standards.ieee.org/standard/7000-2021.html>

Inženierijas lēmumus nosaka pamatnostādne s un metodes

Uz vērtībām balstīts dizains izmanto konceptuālus, empīriskus un tehniskus pētījumus, lai izvērtētu cilvēku vērtības visā dizaina procesā. IEEE 7000 nodrošina sistēmu inženierijas procesu, lai noskaidrotu vērtības un iekļautu tās koncepcijās, prasībās un ētisko risku pārvaldībā.

ES ētikas vadlīnijas nosaka prasības uzticamam mākslīgajam intelektam. ES AI likums nosaka juridiskas saistības noteiktām sistēmām un dalībniekiem. Šie avoti kalpo par vadlīnijām saistītajam darbam un neaizstāj ieinteresēto pušu sniegto informāciju vai nozares inženierijas standartus.

Izmantojiet spēkā esošo tiesību aktu tekstu un reģistrējiet, kāda metode, pienākums un pierādījumi pamato katru nozīmīgo lēmumu.

[Avoti]

- https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf
- <https://standards.ieee.org/standard/7000-2021.html>
- <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02024R1689-20260727>

Pašreizējie pienākumi un inženieru atbildība

ES AI likums nosaka pienākumus atkarībā no sistēmas, riska kategorijas un dalībnieka lomas. Inženieru komandām šie pienākumi jāsaista ar nozares tiesību aktiem, drošības standartiem un organizatoriskajām kontroles procedūrām.

Par katru nozīmīgu lēmumu jāreģistrē: lēmuma pieņēmējs, pierādījumu turētājs, pārbaudītājs, kontroles iestāde un apturēšanas iestāde. Atbildība kļūst praktiski īstenojama, ja pirms ieviešanas ir redzami pienākumi, pierādījumi, apstiprināšanas un eskalācijas ceļi.

[Avoti]

- <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02024R1689-20260727>
- <https://www.nist.gov/itl/ai-risk-management-framework>
- Jaunākā 5. moduļa prezentācija, ko nodrošinājis projekta konsorcijs „Ethical Engineer”.

01

Iepazīstieties ar ieinteresētajām pusēm un vērtībām

Definējiet vērtības kontekstā, izvērtējiet ietekmi un vāciet pierādījumus no iesaistītajām pusēm.

Uz vērtībām balstīta projektēšana saista ētisko mērķi ar sistēmas izvēlēm

Vērtība izsaka to, ko cilvēki vai iestādes uzskata par svarīgu, piemēram, drošību, cieņu, autonomiju, taisnīgumu, privātumu, pieejamību, uzticamību vai ilgtspējību. Tās nozīme ir atkarīga no cilvēkiem, prakses un darbības konteksta.

Vērtības ietekmē mērķus, datus, funkcijas, saskarnes, noklusējumus, biznesa noteikumus un uzturēšanu. Tās var papildināt viena otru vai radīt spriedzi.

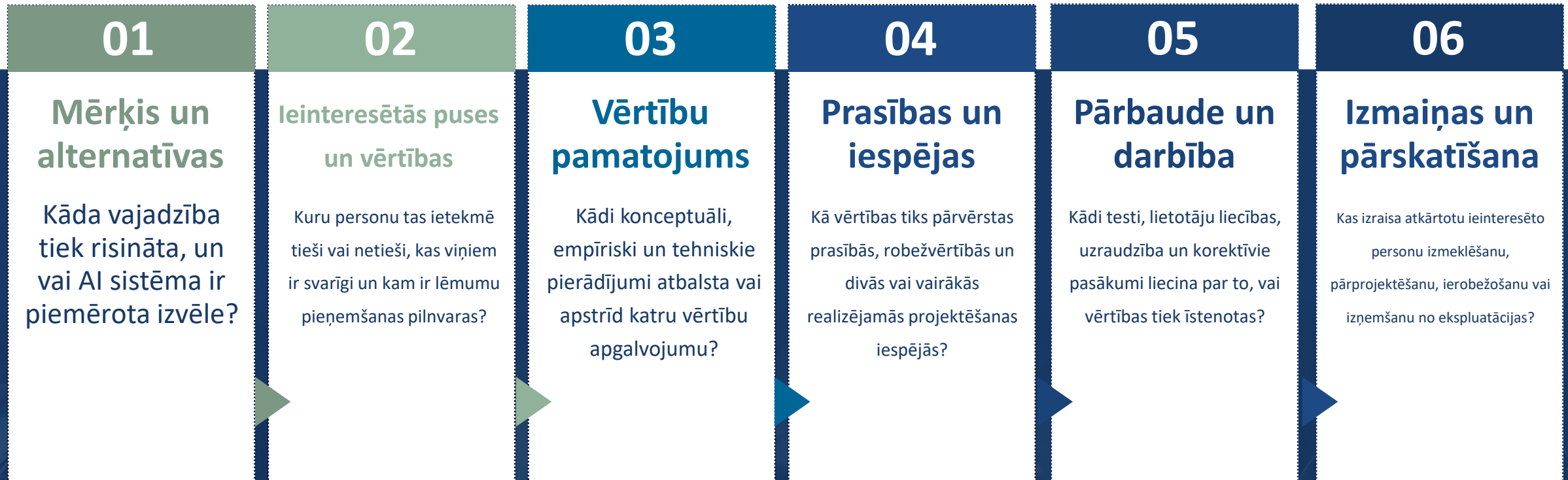
Uz vērtībām balstīts lēmums identificē ieinteresētās puses, pierādījumus un alternatīvas, pēc tam izvēlēto vērtību interpretāciju attēlo prasībās, testos, atbildīgajos un pārskatīšanas nosacījumos.

[Avoti]

- https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf
- <https://standards.ieee.org/standard/7000-2021.html>

Vērtību kartēšana visā inženierijas dzīves ciklā

Izvēlieties vienu inženierijas lēmumu. Katrā posmā identificējiet ieinteresētās puses, vērtību apgalvojumus, pierādījumus, projektēšanas risinājumu un pārskatīšanu.



REZULTĀTS Vērtību karte, kas saista katru posmu ar ieinteresētajām pusēm, pierādījumiem, projektēšanas risinājumu, pieņemšanas kritērijiem un pārskatīšanas iemodu.

Tiešajām un netiešajām ieinteresētajām personām ir atšķirīga ietekme

- Tiešie ieinteresētie lietotāji izmanto, ekspluatē, uztur sistēmu vai saņem no tās lēmumus. Netiešie ieinteresētie lietotāji var tikt ietekmēti, pat nepieskaroties sistēmai, tostarp ģimenes locekļi, kopienas, apkārtējie cilvēki, nākamie darbinieki, piegādātāji un cilvēki, kas pārstāvēti apmācības datos.
- Uzskaitiet, kas iegūst iespējas, kas uzņemas risku, kas sniedz darbaspēku vai datus un kas var tikt izslēgts. Iekļaujiet cilvēkus, kurus ietekmē paredzama launprātīga izmantošana un sistēmas kļūmes.
- Ieinteresēto pušu karte ir nepilnīga, ja tajā ir iekļauti tikai pircēji, lietotāji un regulatori.

Vara nosaka, kuru vērtības kļūst par prasībām

- leinteresētās puses atšķiras pēc lēmumu pieņemšanas pilnvarām, tehniskajām zināšanām, laika, valodas, resursiem un spējas atteikties. Pircēja vadīta apspriešanās var apklusināt cilvēkus, kurus ietekmē šīs sekas.
- Reģistrējiet, kurš formulē problēmu, atlasa pierādījumus, apstiprina prasības un var apstrīdēt ieviešanu. Nodrošiniet neatkarīgas līdzdalības iespējas gadījumos, kad nodarbinātības, aprūpes, sabiedrisko pakalpojumu vai drošības attiecības ierobežo brīvu viedokļu atšķirību.
- Neatnākšanu, vāju atsaucību un nepiekrīšanu uzskatiet par pierādījumu, kas jāizmeklē, nevis par piekrišanu.

Definējiet katru vērtību konkrētā inženierijas kontekstā

Drošība var nozīmēt traumu novēršanu, droša stāvokļa uzturēšanu vai savlaicīgas iejaukšanās atbalstīšanu. Autonomija var attiekties uz apzinātu izvēli, kontroli, brīvību no manipulācijām vai praktisko spēju atteikties. Taisnīgums var attiekties uz rezultātiem, procesu, piekļuvi vai tiesiskās aizsardzības līdzekļiem.

Izstrādājiet kontekstuālu definīciju, norādiet iesaistītās puses, ticamu apdraudējumu un novērojamus pierādījumus. Izvairieties no plaša nozīmes vārdu lietošanas, it kā visiem būtu viena interpretācija.

Definīcijai jābūt pietiekami precīzai, lai tā kalpotu par vadlīniju projektēšanas variantam un pieņemšanas testam.

leinteresēto pušu pierādījumiem ir jāietekmē dizaina izvēles

Atkarībā no lēmuma izmantojiet intervijas, novērojumus, kopīgu projektēšanu, aptaujas, lietojamības pētījumus, sūdzības un pierādījumus no prakses.

Iekļaujiet ietekmētos cilvēkus, gala lietotājus, kā arī tehniskos, juridiskos, ētiskos un nozares ekspertus. Eksperti neaizstāj tos, kurus ietekme skar.

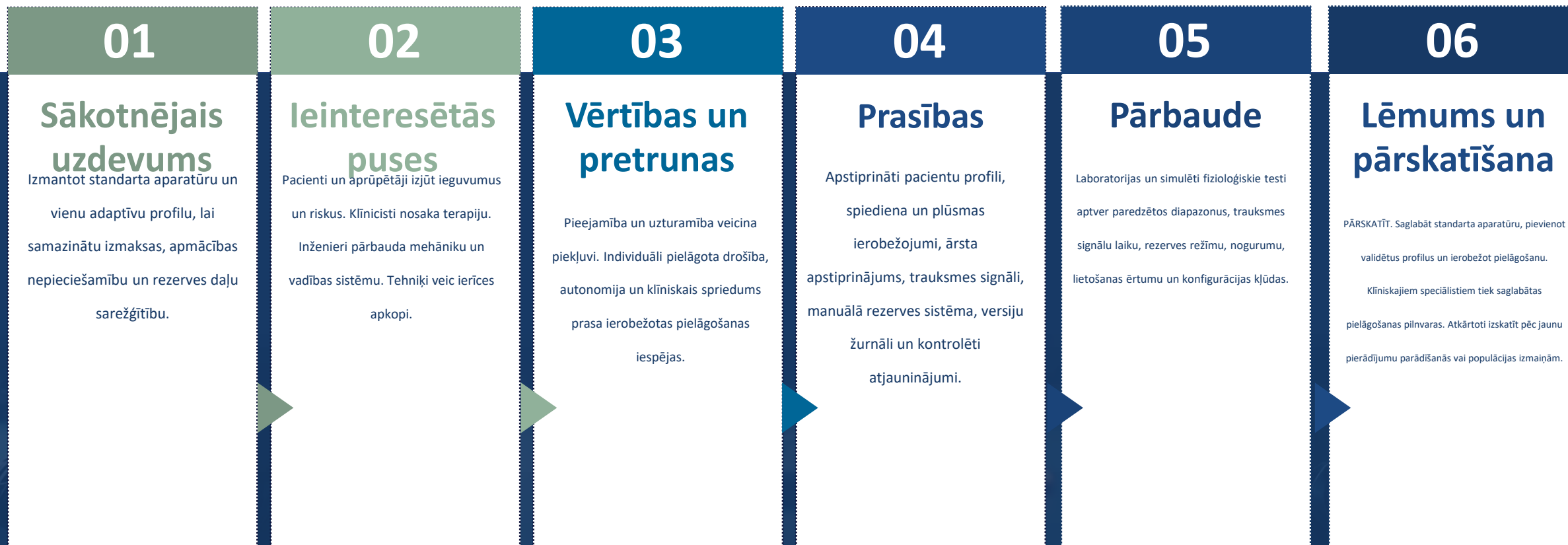
Reģistrējiet, kas ir mainījies: problēmas definīcija, prasība, variants, testa vai pārskatīšanas nosacījums — un kurš joprojām nav pārstāvēts.

Vērtību pretrunas prasa pierādījumus un ierobežojumus

- Vērtības var savstarpēji papildināt viena otru: saprotami paskaidrojumi var uzlabot autonomiju, drošību un uzticēšanos. Tās var arī būt pretrunā: detalizēta datu reģistrēšana var veicināt atbildību, vienlaikus palielinot privātuma risku; personalizācija var uzlabot atbilstību, vienlaikus samazinot standartizāciju un uzturēšanas iespējas.
- Vispirms definējiet neapstrīdamās juridiskās, drošības vai tiesību robežvērtības. Pēc tam salīdziniet iespējamās alternatīvas, izmantojot atbilstošus pierādījumus. Neaprakstiet novēršamu kaitējumu kā līdzsvarotu kompromisu.
- Reģistrējiet atlikušo pretrunu, ietekmētās ieinteresētās puses un pierādījumus, kas varētu likt pārskatīt lēmumu.

PRAKTISKS MEHĀNISKĀS INŽENIERIJAS PIEMĒRS: SIRDS SŪKŅA PROJEKTĒŠANA

Komanda ierosina vienu ar mākslīgo intelektu atbalstītu vadības profilu pulsējošai sirds sūknim, lai vienkāršotu ražošanu un uzturēšanu. Iesaistīto pušu sniegtie pierādījumi liecina, ka dažādiem pacientiem ir nepieciešami apstiprināti darbības diapazoni un skaidra klīniskā pilnvara.



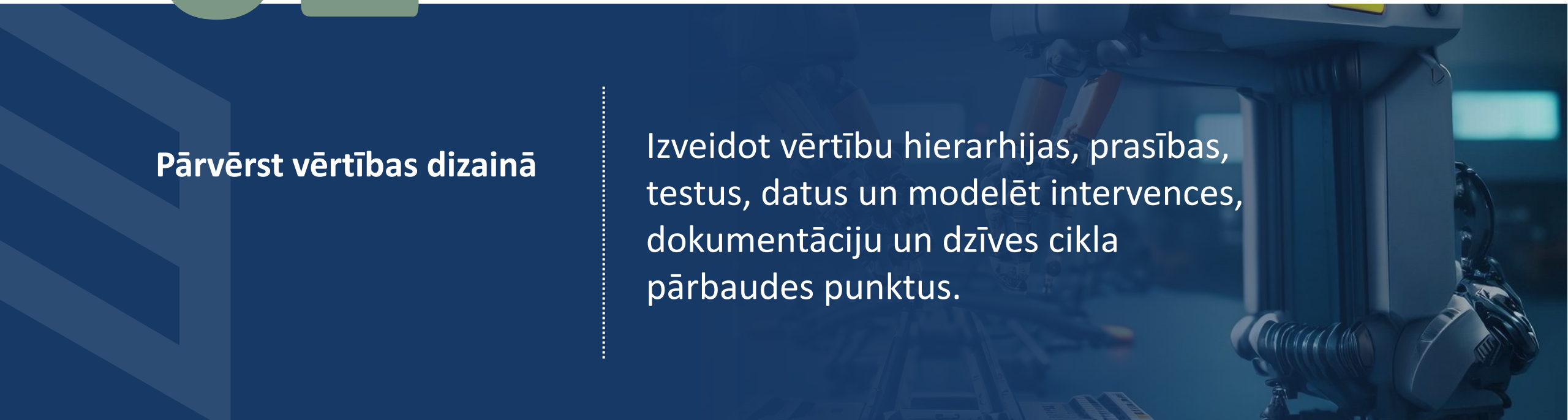
VĒRTĪBAS IETEKMĒJA PROJEKTU Standartizācijas uzdevums pārvērtās par konfigurējamu mehānisko un vadības arhitektūru ar pacientam specifiskiem ierobežojumiem, klīnisko kompetenci un pārskatāmiem pierādījumiem



02

Pārvērst vērtības dizainā

Izveidot vērtību hierarhijas, prasības, testus, datus un modelēt intervences, dokumentāciju un dzīves cikla pārbaudes punktus.



Trīs izpētes veido vienu iteratīvu dizaina metodi

Uz vērtībām orientēts dizains apvieno konceptuālus, empīriskus un tehniskus pētījumus. Katrs no tiem var ietekmēt pārējos, mainoties problēmai, pierādījumiem un dizainam.

Izmantojiet šos trīs pētījumu veidus atkārtoti. Iesaistītās puses secinājums var mainīt prasību; prototips var atklāt jaunu iesaistīto pusi vai vērtību pretrunu.

Skatīt nākamo slaidu...

[Avoti]

- https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf
- <https://standards.ieee.org/standard/7000-2021.html>

Trīs izmeklēšanas veido vienu iteratīvo dizaina metodi

01

Konceptuālais

Identificējiet tiešās un netiešās ieinteresētās puses, definējiet vērtības kontekstā un izvērtējiet pretrunas, tiesības un pieņēmumus.

02

Empīriskais

Izpētīt, kā cilvēki saprot, uztver un rīkojas saistībā ar tehnoloģiju, izmantojot piemērotas sociālās un lauka pētījumu metodes.

03

Tehniskais

Analizēt, kā esošie un alternatīvie risinājumi atbalsta vai vājina vērtības, pēc tam izstrādāt prototipu un pārbaudīt prasības.

Rezultāts ir izsekojama ķēde: vērtību definīcija, ieinteresēto pušu pierādījumi, dizaina prasības, tehniskā atbilde, pieņemšanas tests, īpašnieks un pārskatīšanas iemesls.

Vērtību hierarhija padara tulkojumu pārskatāmu

- Sāciet ar kontekstuālu vērtību, pēc tam identificējiet to atbalstošās normas vai projektēšanas principus. Pārtulkojiet katru normu vienā vai vairākās prasībās un izmērāmos pieņemšanas kritērijos.
- Piemērs: darbinieka autonomija → uzraudzība attiecas tikai uz iekārtu drošību → individuālie ražīguma rādītāji ir aizliegti → pārliecinieties, ka saskarne, datu bāze un ziņojumi tos nevar ģenerēt.
- Reģistrējiet pierādījumus, atbildīgo personu un statusu katrai saitei. Ja prasību nevar saistīt ar vērtību un ieinteresēto personu vajadzībām, apšaubiet tās prioritāti.

Ētiskajiem lietotāju stāstiem nepieciešami pārbaudāmi pieņemamības kritēriji

- Noderīgs stāsts identificē ietekmēto lomu, spēju vai aizsardzību un iemeslu, kāpēc tas ir svarīgi. Tajā tiek norādīti arī ļaunprātīgas izmantošanas, izslēgšanas vai kļūdu apstākļi.
- Piemērs: Kā apkopes tehniķim man ir nepieciešams, lai sistēma parādītu sensora uzticamību un nesenās konfigurācijas izmaiņas, lai es varētu novērtēt, vai brīdinājums ir rīcībai nepieciešams.
- Pieņemamības pierādījumi: reprezentatīvs uzdevuma tests, skaidri formulēta nenoteiktība, versiju vēsture, brīdinājuma reaģēšanas laiks un dokumentēts rezerves risinājums. Saistiet stāstu ar prasību un verifikācijas ierakstu.

Datu intervencēm ir saistītas vērtību izvēles

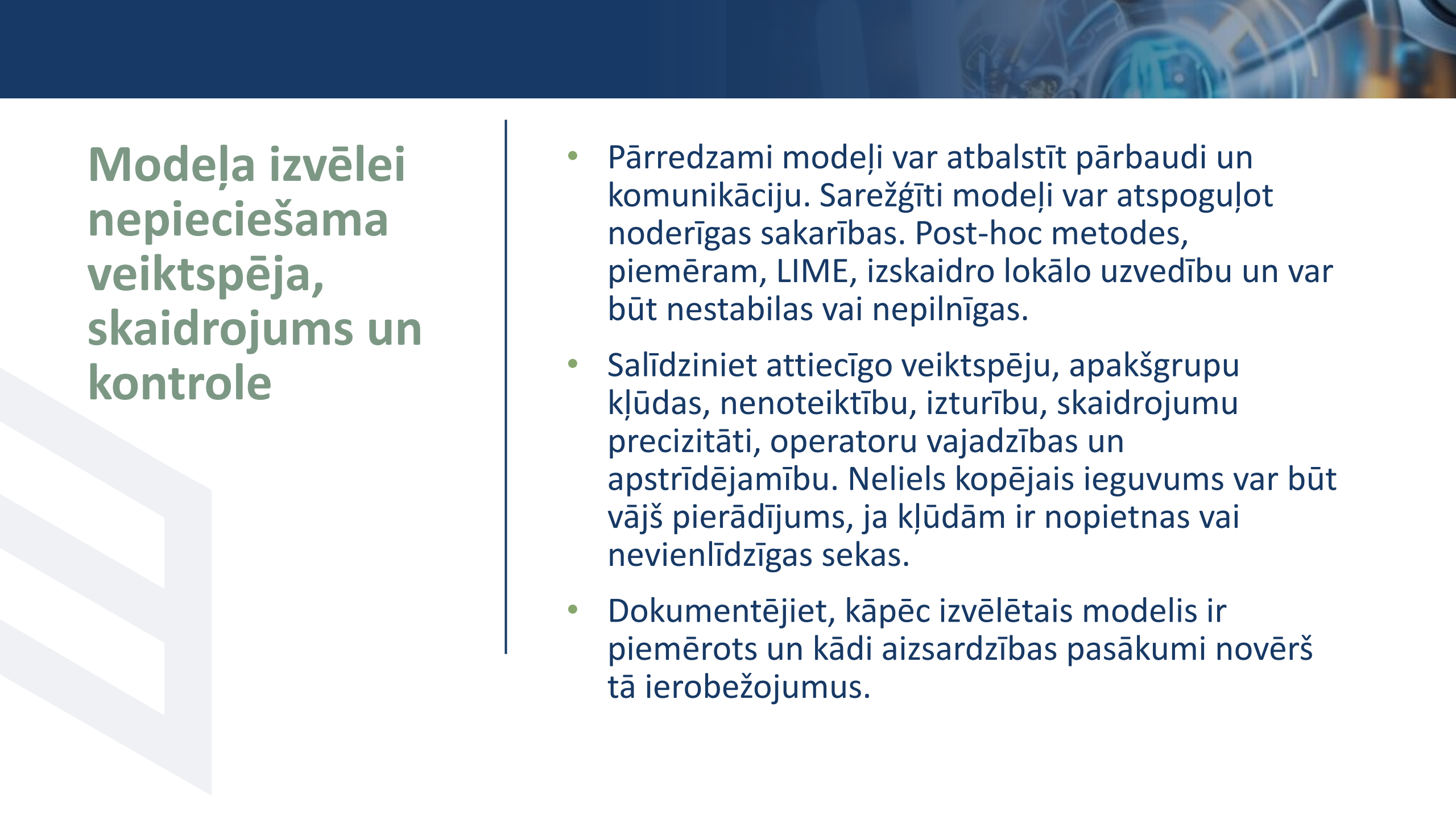
- Paraugu ņemšana, marķējumi, aizstājēji, trūkstošo datu noteikumi un papildināšana nosaka, kuriem gadījumiem tiek pievērsta uzmanība. SMOTE vai klases svēršana var atbalstīt retas klases, vienlaikus mainot kalibrēšanu. Pārnesamā mācīšanās pārņem pieņēmumus no citas jomas. PCA vai pazīmju atlase var uzlabot pārskatāmību, vienlaikus noņemot informāciju, kas attiecas uz apakšgrupu.
- Katrai intervencijai norādiet paredzēto vērtību, ietekmēto grupu, mehānismu, salīdzinājuma bāzes līniju un novērtēšanas pierādījumus.
- Dokumentējiet izcelsmi, ierobežojumus, aizliegtos lietojumus un lēmumu saglabāt vai noņemt iezīmi.

Datu nelīdzsvarotība un SMOTE prasa rūpīgu novērtēšanu

Retās klases var atspoguļot paraugu ņemšanas, marķēšanas, piekļuves vai mērīšanas problēmas. Pirms izvēlaties tehnisko risinājumu, izpētiet cēloni.

Iespējas ietver jaunu datu vākšanu, stratificētu novērtēšanu, klašu svēršanu un sintētisko mazākuma pārmērīgu paraugu ņemšanu. SMOTE rada sintētiskus mazākuma piemērus starp novērotajiem gadījumiem; tas var atbalstīt mācīšanos, vienlaikus mainot kalibrēšanu vai radot neiespējamās situācijas, ja pazīmju telpa nav piemērota.

Salīdziniet bāzes līniju un intervenci, izmantojot apakšgrupu kļūdas, kalibrēšanu, nenoteiktību un domēna pārskatu. Reģistrējiet, kādu vērtību intervence atbalsta un kādi riski paliek.



Modeļa izvēlei nepieciešama veiktspēja, skaidrojums un kontrolē

- Pārredzami modeļi var atbalstīt pārbaudi un komunikāciju. Sarežģīti modeļi var atspoguļot noderīgas sakarības. Post-hoc metodes, piemēram, LIME, izskaidro lokālo uzvedību un var būt nestabilas vai nepilnīgas.
- Salīdziniet attiecīgo veiktspēju, apakšgrupu kļūdas, nenoteiktību, izturību, skaidrojumu precizitāti, operatoru vajadzības un apstrīdējamību. Neliels kopējais ieguvums var būt vājš pierādījums, ja kļūdām ir nopietnas vai nevienlīdzīgas sekas.
- Dokumentējiet, kāpēc izvēlētais modelis ir piemērots un kādi aizsardzības pasākumi novērš tā ierobežojumus.

Interpretējamība un LIME sniedz lokālus pierādījumus

LIME piemēro vienkāršu lokālu aizstājēju ap vienu prognozi, lai novērtētu, kuras pazīmes ietekmēja šo lokālo rezultātu.

Tas var palīdzēt inženieriem izpētīt uzvedību, apspriest konkrētu gadījumu un identificēt jautājumus turpmākai pārbaudei. Paskaidrojums var mainīties atkarībā no paraugu ņemšanas, parametriem un korelētajām pazīmēm.

LIME nenodrošina cēloņsakarību, globālo uzvedību vai taisnīgumu. Pārbaudiet stabilitāti atkārtotos izpildes ciklos un līdzīgos gadījumos, salīdziniet ar nozares zināšanām un reģistrējiet lēmumu, ko pamato izskaidrojums, kā arī ierobežojumus, par kuriem tiek informēti lietotāji.

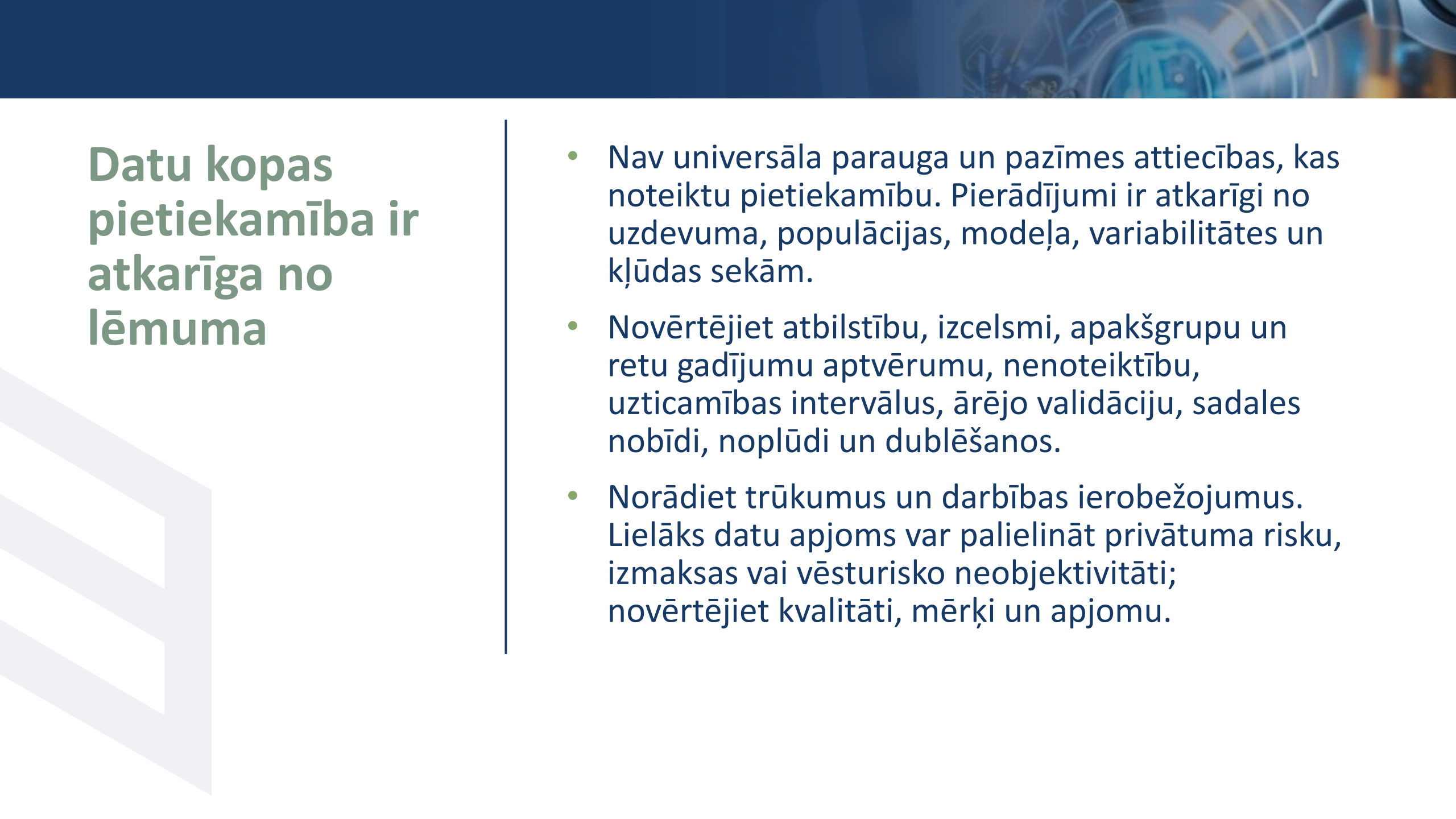
Saskarnes veido uzmanību, autoritāti un izaicinājumus

Noklusējumi, brīdinājumu secība, pārliecības rādīšana, apzīmējumi un darba plūsmas laiks ietekmē uzticēšanos.

Automatizācija maina cilvēka lomu. Kad operatori kļūst par uzraugiem, uzticama automatizācija var samazināt modrību un palēnināt iejaukšanos.

Pārbaudiet, vai lietotāji spēj atpazīt ierobežojumus, atgūt situācijas izpratni, rīkoties laicīgi, pārrakstīt lēmumu, lūgt pārskatīšanu un atgūties pēc kļūdas.

Reģistrējiet informāciju, kas nepieciešama katrai mērķauditorijai, un darbības, kuras saskarnei ir jāatbalsta.



Datu kopas pietiekamība ir atkarīga no lēmuma

- Nav universāla parauga un pazīmes attiecības, kas noteiktu pietiekamību. Pierādījumi ir atkarīgi no uzdevuma, populācijas, modeļa, variabilitātes un kļūdas sekām.
- Novērtējiet atbilstību, izcelsmi, apakšgrupu un retu gadījumu aptvērumu, nenoteiktību, uzticamības intervālus, ārējo validāciju, sadales nobīdi, noplūdi un dublēšanos.
- Norādiet trūkumus un darbības ierobežojumus. Lielāks datu apjoms var palielināt privātuma risku, izmaksas vai vēsturisko neobjektivitāti; novērtējiet kvalitāti, mērķi un apjomu.

Datu apjoms, dimensiju skaits un pārnesamā mācīšanās

Nav universāla parauga noteikuma. Pietiekamība ir atkarīga no uzdevuma, populācijas, variabilitātes, modeļa un kļūdu sekām.

Augstas dimensijas dati var radīt retumu un pārmācīšanos pat tad, ja kopējais paraugs šķiet liels. Izmantojiet mācīšanās līknes, pazīmju pārskatīšanu, regularizāciju, nenoteiktības novērtējumus un ārējo validāciju.

Pārnesamā mācīšanās un datu papildināšana var samazināt datu prasības, vienlaikus pārņemot pieņēmumus no citas jomas. Pārbaudiet atbilstību jomai, apakšgrupu pārklājumu un sadales nobīdi.

Dokumentācija saista dizainu, pierādījumus un pārskatīšanu

- Datu kopas datu lapā tiek reģistrēts pamatojums, sastāvs, vākšana, apstrāde, ieteicamie lietojumi un ierobežojumi. Modeļa kartē tiek reģistrēts paredzētais lietojums, novērtējums, apakšgrupu rezultāti, ierobežojumi un ētiskie apsvērumi. Sistēmas kartē tiek savienoti komponenti, saskarnes, lietotāji, darbības konteksts un aizsardzības pasākumi.
- Izmantojiet dokumentāciju projektēšanas pārskatīšanas, iepirkumu, izmaiņu kontroles un incidentu analīzes laikā. Uzturiet to versijās un saistiet ar pierādījumiem.
- Dokumentācija veicina atbildību, ja tā maina lēmumu, atklāj nepilnības vai ļauj veikt pārskatīšanu.

Datu lapas, modeļu kartes un sistēmu kartes

Datu kopu datu lapās tiek reģistrēti pamatojums, sastāvs, vākšana, apstrāde, paredzētais lietojums un ierobežojumi.

Modeļu kartēs tiek reģistrēts paredzētais lietojums, novērtējums, apakšgrupu rezultāti, ierobežojumi un ētiskie apsvērumi. Sistēmas kartēs tiek savienoti komponenti, saskarnes, lietotāji, konteksts un aizsardzības pasākumi.

Saglabāji ierakstus ar versiju numuriem un saistiet tos ar pierādījumiem, apstiprinājumiem un izmaiņu kontroli. Ja ir piemērojams ES AI likums, saistiet tos ar tehnisko dokumentāciju, riska pārvaldību un pēcpārdošanas uzraudzības pierādījumiem.

Ātrās pārbaudes nodrošina vērtību saglabāšanu piegādes darba plūsmā

- Pievienojiet vērtības jautājumus atklāšanai, uzkrāto uzdevumu precizēšanai, dizaina pārskatīšanai, sprinta plānošanai, pabeigšanas definīcijai, izlaišanai un retrospekcijai. Saistiet katru ētisko lietotāja stāstu ar prasību, testu un pierādījumu īpašnieku.
- Pārbaudes punktam jābūt spējīgam aizkavēt vai mainīt darbu. Reģistrējiet neatrisinātās vērtību pretrunas, ieinteresēto pušu pierādījumu trūkumus un nosacījumus ierobežotai izlaišanai.
- Ja mainās dati, modelis, saskarne, piegādātājs vai darbības konteksts, jānosaka, kuras vērtību prasības un ieinteresēto pušu pierādījumi ir jāpārskata no jauna.

Uzraudzībai jāatklāj gan vērtību novirzes, gan modeļu novirzes

- Tehniskā veikspēja var palikt stabila, kamēr mainās sociālais konteksts. Jauna mērķauditorija, darba prakse, stimulsi, politika vai sekundāra izmantošana var apdraudēt sākotnējo vērtības piedāvājumu.
- Uzraugiet attiecīgos rezultātus, sūdzības, pārrakstījumus, izņēmumus, apvedceļus, incidentus un sadales ietekmi. Nodrošiniet pieejamu atgriezeniskās saites un problēmu novēršanas ceļu.
- Katram rādītājam noteikt atbildīgo personu, pārskatīšanas biežumu, sliekšni un veicamos pasākumus. Atkārtoti izvērtēt ieinteresētās puses un prasības, ja pierādījumi pārsniedz noteikto sliekšni.

Pirms svērto variantu vērtēšanas piemērojiet obligātos sliekšņus

Sāciet ar juridiskajiem, drošības, tiesību un profesionālajiem ierobežojumiem. Izslēdziet jebkuru variantu, kas neatbilst obligātajam sliekšnim, ja vien dizainu nav iespējams pārskatīt.

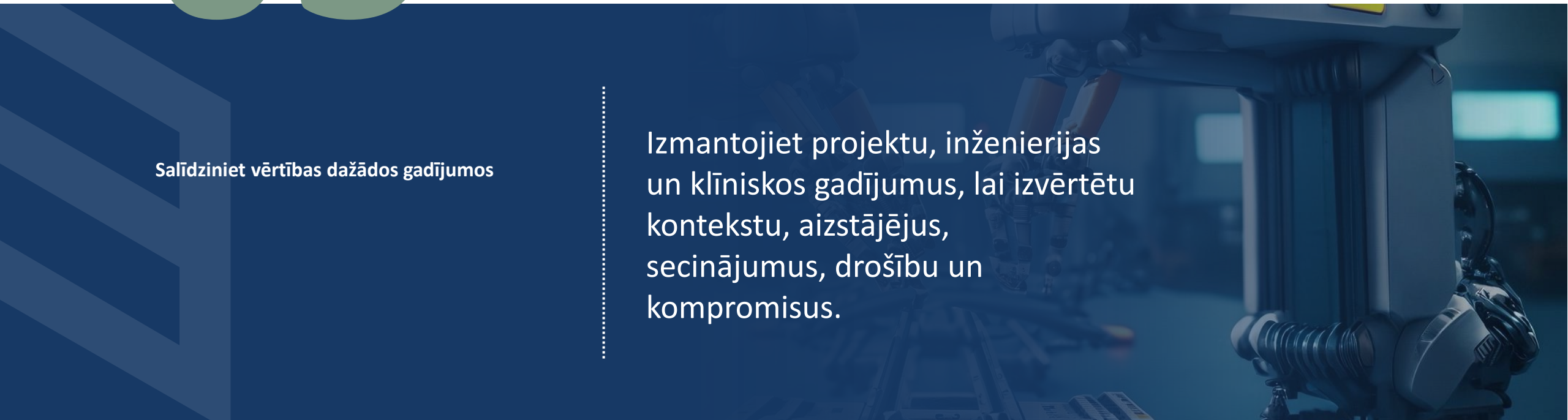
Saistiet uz vērtībām balstītu dizainu ar riska pārvaldību: identificējiet kaitējumu vai draudus vērtībām, novērtējiet pierādījumus un sekas, izvēlieties dizaina kontroles pasākumu, pārbaudiet kontroles pasākumu un dokumentējiet atlikušo risku, kā arī pārskatiet sliekšņus.

Tad salīdziniet iespējamās opcijas, izmantojot tādas kritērijus kā drošība, autonomija, taisnīgums, privātums, izskaidrojamība, veiktspēja, ilgtspēja, izmaksas un uzturamība. Svērtais koprezultāts nevar atcelt nopietnu tiesību vai drošības pārkāpumu.

03

Salīdziniet vērtības dažādos gadījumos

Izmantojiet projektu, inženierijas un klīniskos gadījumus, lai izvērtētu kontekstu, aizstājējus, secinājumus, drošību un kompromisus.



Projekta gadījums: AI un darbinieku atlase

- „Ethical Engineer” darbinieku atlases gadījums izskata, kā automatizācija var ietekmēt piekļuvi darbam, reproducēt vēsturiskos modeļus un mainīt lēmumu pieņemšanas varu starp kandidātiem, personāla atlases speciālistiem un piegādātājiem.
- Inženierijas fokuss: mērķis, pierādījumi, būtiskās īpašības, pieejamība, paskaidrojums, personāla atlases speciālista autoritāte, izaicinājums un uzraudzība.
- Atveriet projekta gadījumu par darbinieku atlasī. Identificējiet vienu organizācijas izvirzīto vērtību apgalvojumu un pierādījumus, kas nepieciešami, lai to pārbaudītu.

Projekta gadījums: vēsturisko fotogrāfiju novērtēšana

- „Ethical Engineer” gadījums pēta AI atbalstītu vēsturisko fotogrāfiju novērtēšanu. Kultūras nozīmi, izcelsmi un nenoteiktību var samazināt, ja modelis attēlu uztver kā atsevišķus datus.
- Inženierijas fokuss: avota konteksts, pārstāvētās kopienas, jutīgi secinājumi, ekspertu interpretācija, nenoteiktība, piekļuve un labojumi.
- Atveriet vēsturisko fotogrāfiju gadījumu. Noteikt, kurām ieinteresētajām pusēm būtu jāietekmē paredzētais lietojums un kādi secinājumi būtu jāaizliedz vai jāierobežo.

Projekta gadījums: AI datu analītikā

- „Ethical Engineer” datu analīzes gadījumā tiek izskatīts, kā tehniski pamatotai analīzei var rasties jaunas sekas, ja dati tiek apvienoti, no tiem tiek izdarīti secinājumi, tie tiek vizualizēti vai atkārtoti izmantoti citā lēmumā.
- Inženierijas fokuss: mērķis, izcelsme, aizstājēji, apkopojums, interpretācija, piekļuve, sekundārā izmantošana un tas, kurš var apstrīdēt secinājumu.
- Atveriet projekta datu analīzes gadījumu. Identificējiet vērtību pretrunu, ko rada viena datu vai vizualizācijas izvēle, un ierosiniet pārbaudāmu dizaina risinājumu.

Projekta gadījums: AI pulsējošās sirds sūkņos

- „Ethical Engineer” sirds sūkņu gadījums izskata standartizāciju, piekļuvi, personalizētu aprūpi un drošību. Dizains, kas vienkāršo ražošanu un apkopi, var neņemt vērā fizioloģiskās atšķirības vai klīnisko vērtējumu.
- Inženierijas fokuss: mērķa populācija, pacientu vērtības, drošas mehāniskās un kontroles robežas, klīniskā personāla autoritāte, trauksmes signāli, pierādījumi un atbildība par pielāgošanu.
- Atveriet pulsējošo sirds sūkņu gadījumu. Salīdziniet fiksētu, konfigurējamu un adaptīvu dizainu, izmantojot obligātos drošības sliekšņus.

Tā pati metode atklāj dažādas vērtību pretrunas

- Darbinieku atlase uzsver taisnīgu piekļuvi, cieņu, paskaidrojumus un izaicinājumus. Vēsturiskās fotogrāfijas piešķir kultūras kontekstu, izcelsmi un cieņpilnu interpretāciju. Datu analītika izceļ mērķi, secinājumus, aizstājēju izvēli un sekundāro izmantošanu. Sirds sūkņi iekļauj drošību, piekļuvi, personalizāciju un klīnisko autoritāti mehāniskajā sistēmā.
- Katram gadījumam definējiet inženiertehnisko lēmumu, ieinteresētās puses, konteksta vērtības, pierādījumu trūkumu, obligāto sliekšni un iespējamās opcijas.
- Salīdziniet, kuras metodes ir pārnesamas un kuras prasības ir atkarīgas no jomas.

Izmaksu ierobežojumi nevar atcelt obligātos drošības sliekšņus

- Redundance, uzraudzība un verifikācija palielina izmaksas, svaru, enerģijas patēriņu, apmācību vai apkopi. Šīs sekas ir jāizvērtē, taču budžeta spiediens nedara nedrošu risinājumu pieņemamu.
- Pirms izmaksu salīdzināšanas noteikt minimālo drošības funkciju. Projektēt, ņemot vērā pakāpenisku darbības pasliktināšanos: drošu stāvokli, būtiskas funkcijas, rezerves maršrutu un nosacījumus cilvēka iejaukšanās gadījumam.
- Izpētiet pārprojektēšanu, pakāpenisku darbības jomas paplašināšanu, alternatīvu arhitektūru un risinājumu bez mākslīgā intelekta izmantošanas. Reģistrējiet, kurš sedz katras izmaksas un risku.

Kļūmju gadījumi atklāj vērtības, kas ietvertas iepriekšējos lēmumos

„Uber” parāda, kā grafiks, testu pārvaldība un pieņēmumi par cilvēka veikto uzraudzību var ietekmēt drošību. „Boeing 737 MAX” parāda sensoru atkarības, automatizācijas pilnvaru, informācijas un organizatoriskā spiediena sekas. Klīniskais astmas gadījums parāda, kā modelis var apgūt aprūpes shēmas, nevis fizioloģiju.

Atrodiet iepriekšējo projektēšanas lēmumu, kur labāki pierādījumi, ieinteresēto pušu iebildumi vai obligāta robežvērtība varētu būt mainījuši rezultātu.

Neizmantojiet procedūru kā pārprojektēšanas aizstājēju, ja briesmas slēpjas arhitektūrā. Kļūdu analīzes rezultātā jāizstrādā pārskatītas prasības un verifikācijas plāns.

Klīniskais AI gadījums: astma un aprūpes procesa traucējošie faktori

Šķita, ka slimnīcas riska modelis uzskatīja astmu par aizsargājošu faktoru pacientiem ar pneimoniju. Apmācību kontekstā pacienti ar astmu bieži vien agrāk saņēma intensīvo aprūpi. Modelis apguva aprūpes gaitu un tās rezultātus, nevis drošu fizioloģisku likumsakarību.

Šis raksts var kļūt nedrošs, ja to izmanto vidē ar atšķirīgām uzņemšanas praksēm. Inženieriem jāizvērtē cēloņsakarības pieņēmumi, aprūpes procesi, pārnesamība un jomas maiņa.

Pārskatiet prasību: klīniskām rekomendācijām ar lielu ietekmi pirms lietošanas ir nepieciešama jomas pārskatīšana, aprūpes ceļa analīze un validācija paredzētajā vidē.

Klīniskais AI gadījums: veselības aprūpes izmaksas kā aizstājējs

Tika izmantots iedzīvotāju veselības algoritms, kurā prognozētās veselības aprūpes izmaksas tika izmantotas kā veselības vajadzību aizstājējs. Nevienlīdzīgā piekļuve un izdevumi nozīmēja, ka izmaksas neatspoguļoja slimības smagumu vienādi visās grupās. Cilvēki ar līdzīgām veselības vajadzībām varēja saņemt atšķirīgus riska novērtējumus un atšķirīgu piekļuvi atbalstam. Jāpārbauda mērķis, aizstājējs, marķējumi un piešķiršanas process. Jāsalīdzina rezultāti pa attiecīgajām grupām un jājaucā iesaistītajām pusēm, vai mērķis atspoguļo paredzēto vērtību. Jāpārveido mērķis, izmantojot tiešus vajadzību rādītājus, jāpārbauda piešķiršanas efekts un jāreģistrē atlikusī nenoteiktība.

Tehniskie pasākumi vienlaikus maina vairākas vērtības

- SMOTE vai klases svēršana var uzlabot uzmanību retām klasēm, vienlaikus ietekmējot kalibrēšanu. Pārnesamā mācīšanās var samazināt datu apjomu un aprēķinu slodzi, vienlaikus importējot neatbilstošas reprezentācijas. Datu papildināšana var uzlabot izturību vai radīt nereālistiskus gadījumus. PCA var samazināt sarežģītību, vienlaikus slēpjot pazīmes, kas ir būtiskas apakšgrupai. LIME var palīdzēt sniegt lokālus paskaidrojumus, vienlaikus saglabājot aptuvenu raksturu.
- Salīdziniet katru intervenci ar bāzes līniju. Pārbaudiet paredzēto labumu, blakusparādības, apakšgrupu sniegumu, nenoteiktību un darbības ierobežojumus.
- Šis rīks ir projektēšanas iespēja, nevis pierādījums tam, ka vērtība ir sasniegta.

Piemērojamais uzdevums:

Veiciet strukturētu vērtību izpēti

1. Izvēlieties darbinieku atlasi, vēsturiskas fotogrāfijas, datu analītiku vai sirds sūkņus.
2. Noteikt inženiertehnisko lēmumu un alternatīvu bez mākslīgā intelekta izmantošanas.
3. Kartējiet tiešās un netiešās ieinteresētās puses, varu un iespējamo izslēgšanu.
4. Noteikt divas kontekstam atbilstošas vērtības un identificēt pretrunu.
5. Norādiet konceptuālos, empīriskos un tehniskos pierādījumus.
6. Izveidojiet divas realizējamās projektēšanas iespējas un obligātos sliekšņus.
7. Izvēlieties variantu, prasību, pieņemšanas testu, atbildīgo personu un pārskatīšanas iemeslu.

Rezultāts: vienas lappuses vērtību izpēte un divu minūšu dizaina aizstāvēšana.

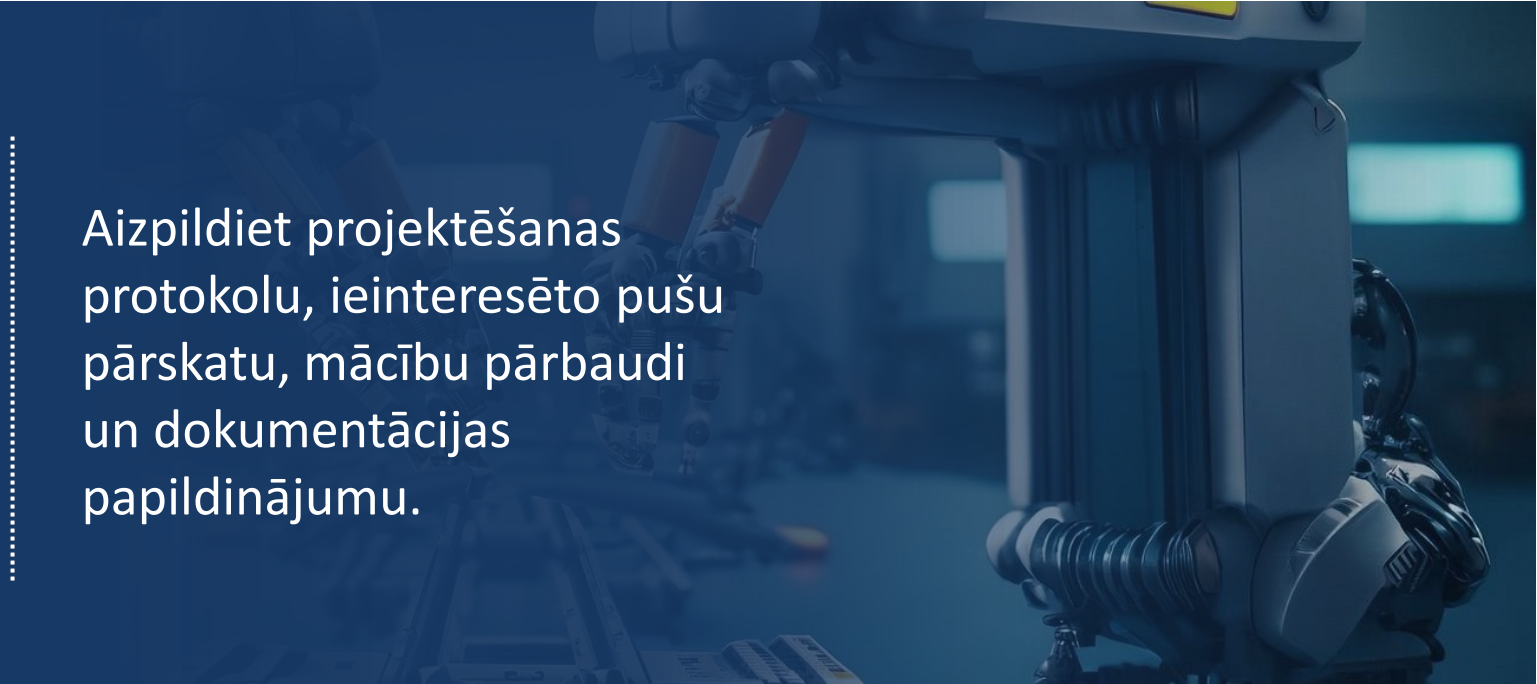
Dilemma: ieinteresēto pušu grupas nevar vienoties par pieņemamu dizainu

- Ar mākslīgo intelektu atbalstīta darba vietas sistēma var samazināt fizisko slodzi. Vadība investīciju lēmumu pieņemšanā novērtē produktivitātes datus. Darbinieki atbalsta drošības brīdinājumus, bet noraida individuālo uzraudzību. Piegādātājs apgalvo, ka sistēma nevar nodalīt šīs divas funkcijas bez papildu izmaksām.
- Nolemjiet: turpināt (GO), pārskatīt (REVISE) vai pārtraukt (STOP). Nosakiet, kādiem pierādījumiem trūkst, kuras vērtības ir pretrunā un kura prasība ir obligāta. Salīdziniet pārprojektēšanu, ierobežotu izmēģinājuma projektu un alternatīvu bez mākslīgā intelekta.
- Reģistrējiet spēcīgāko atšķirīgo viedokli un pierādījumus, kas varētu mainīt lēmumu.

04

Aizstāviet uz vērtībām balstītu lēmumu

Aizpildiet projektēšanas
protokolu, ieinteresēto pušu
pārskatu, mācību pārbaudi
un dokumentācijas
papildinājumu.



Gadījums A: ar AI atbalstīts rūpnieciskais eksoskelets

- Rūpnīca ierosina izmantot ar sensoriem aprīkotu eksoskeletu, lai samazinātu traumu skaitu, uzlabotu pielāgojumu un plānotu apkopi. Piegādātājs piedāvā arī individuālos ražīguma rādītājus, kas balstās uz kustībām, slodzi un dīkstāves laiku. Darbinieki ir nobažījušies par ietekmi uz veselību un disciplināro izmantošanu.
- Salīdziniet trīs variantus: personiskā optimizācija mākonī ar ražīguma analīzi; drošība un pielāgojums ierīcē ar pseidonimizētiem apkopes notikumiem; vai neadaptīva mehāniska sistēma.
- Pirms iespējamo variantu novērtēšanas definējiet obligātās drošības, privātuma un darbinieku līdzdalības prasības.

Pabeigtais uz vērtībām balstītais projektēšanas protokols: rūpnieciskais eksoskelets

LĒMUMS PĀRSKATĪT: Izvēlēties ierīcē integrētu drošības un piemērotības atbalstu; izslēgt individuālo produktivitātes novērtējumu

IEINTERESĒTO PUŠU LIECĪBAS: Darbinieki, drošības pārstāvji, apkopes personāls un uzraugi par prioritāti uzskata piemērotību, drošu slodzes uzlikšanu, skaidrus brīdinājumus un to, ka netiek piemērotas disciplinārās sankcijas.

SLIEKŠŅI: mehāniskie drošības ierobežojumi ir ievēroti; dati netiek izmantoti ražīguma vai apmeklējuma novērtēšanai; darbinieku pārstāvji apstiprina datu robežas; joprojām ir pieejama lietošana saskaņā ar vietējo rokasgrāmatu.

PRASĪBAS UN TESTI: Apstrādāji locītavu leņķus un slodzi lokāli; izdzēsiet neapstrādātos kustības buferus; pseudonimizējiet apkopes notikumus; pārbaudiet slodzes ierobežojumus, brīdinājumu laiku, rezerves risinājumus, lietojamību un ziņojumu saturu.

PĀRSKATĪŠANA: Drošības inženieris ir atbildīgs par slodzes ierobežojumiem; datu vadītājs pārbauda aizliegto izmantošanu; darbinieku pārstāvis pārskata ziņojumus. Atkārtoti jāpārskata pēc funkciju, datu, darba vietas vai piegādātāja izmaiņām.

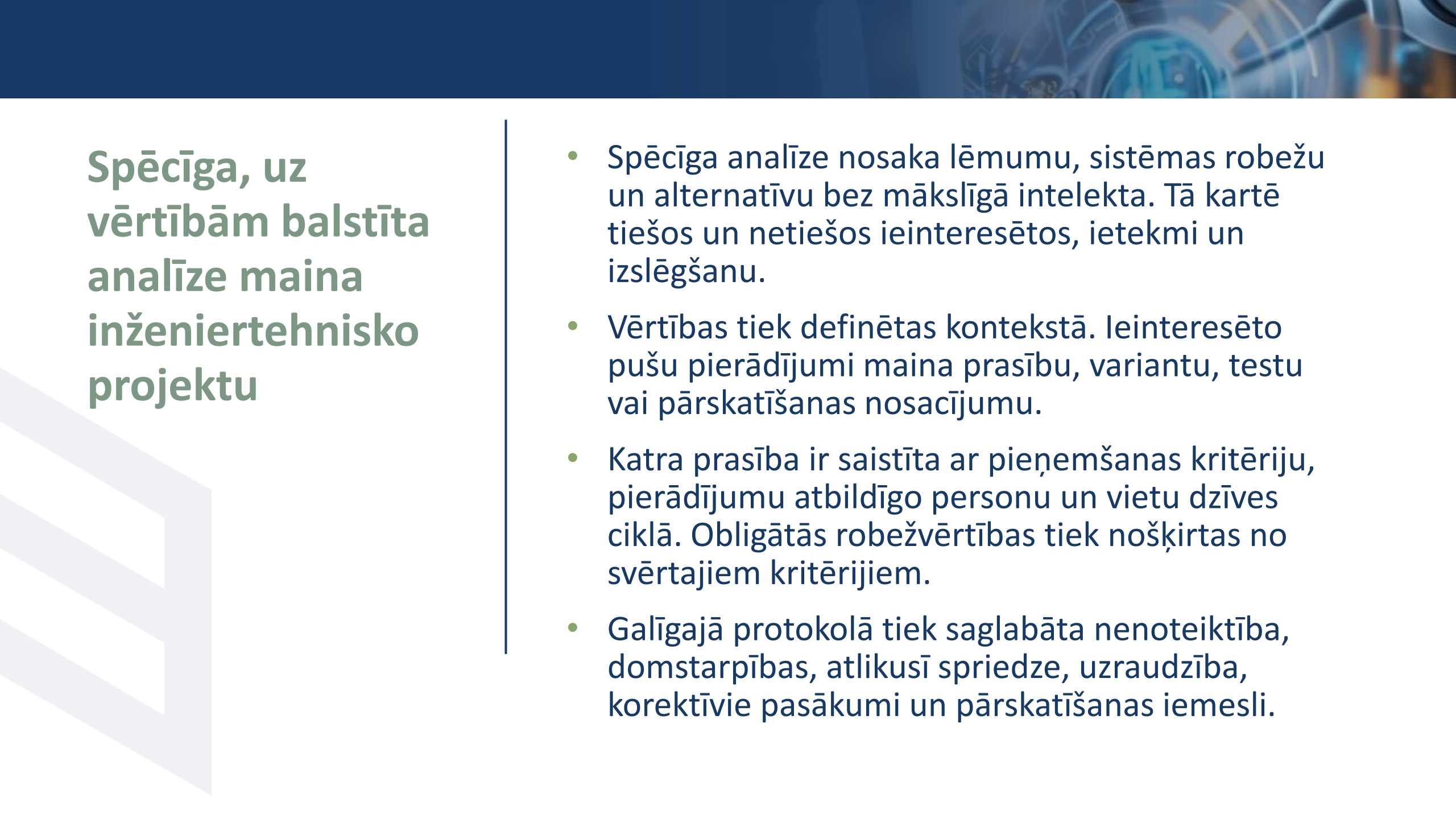
Gadījums B: centralizētās siltumapgādes pieprasījuma kontrole

- Pašvaldība plāno ieviest centrālās apkures piegādes kontroli, izmantojot mākslīgo intelektu, ņemot vērā laika apstākļus, ēku sensorus un pieprasījuma prognozes. Šāds risinājums var samazināt degvielas patēriņu, savukārt konservatīva kontrole var nodrošināt komfortu slikti izolētās mājās. Īrnieki nevar izvēlēties ēkas sensorus, un daži no viņiem nevar iesniegt sūdzības tieši.
- Salīdziniet energoefektivitāti, termisko komfortu, pieejamību, privātumu, noturību un atbildību sabiedrības priekšā. Iekļaujiet uz noteikumiem balstītu un operatora vadītu alternatīvu.
- Pirms ieviešanas noteikt minimālās temperatūras un atjaunošanās sliekšņus, sensoru izmantošanas ierobežojumus, operatora pārrakstīšanas iespējas, publisko pārskatu sniegšanu un sūdzību iesniegšanas kārtību.

Komandas aktivitāte:

Uz vērtībām balstītu
projektēšanas lēmumu
laboratorija

1. Izvēlieties eksoskeletu, centralizēto siltumapgādi vai projekta gadījumu. Norīkojiet sistēmu inženieri, ieinteresēto pušu vadītāju, pierādījumu vadītāju, darbības pārstāvi, kritisko vērtētāju un protokolistu.
2. Noteikt lēmumu, robežu un alternatīvu bez mākslīgā intelekta izmantošanas.
3. Veiciet konceptuālus, empīriskus un tehniskus pētījumus.
4. Pārvērtiet divas vērtības prasībās un pieņemšanas testos.
5. Noteikt obligātās robežvērtības, pēc tam salīdzināt divas vai trīs iespējamās opcijas.
6. Nolemjiet – turpināt (GO), pārskatīt (REVISE) vai pārtraukt (STOP) – un reģistrējiet aizsardzības pasākumus, atbildīgos, atlikušās nesaskaņas un pārskatīšanas nosacījumus.
7. Veiciet divu minūšu prezentāciju un pārskatiet to pēc ieinteresēto pušu iebildumiem.



Spēcīga, uz vērtībām balstīta analīze maina inženiertehnisko projektu

- Spēcīga analīze nosaka lēmumu, sistēmas robežu un alternatīvu bez mākslīgā intelekta. Tā kartē tiešos un netiešos ieinteresētos, ietekmi un izslēgšanu.
- Vērtības tiek definētas kontekstā. Ieinteresēto pušu pierādījumi maina prasību, variantu, testu vai pārskatīšanas nosacījumu.
- Katra prasība ir saistīta ar pieņemšanas kritēriju, pierādījumu atbildīgo personu un vietu dzīves ciklā. Obligātās robežvērtības tiek nošķirtas no svērtajiem kritērijiem.
- Galīgajā protokolā tiek saglabāta nenoteiktība, domstarpības, atlikusī spriedze, uzraudzība, korektīvie pasākumi un pārskatīšanas iemesli.

Darba lapa: Uz vērtībām balstīta projektēšanas dokumentācija

Sistēma, paredzētais lietojums un alternatīva bez mākslīgā intelekta:

Tiešie un netiešie ieinteresētie, vara un izslēgšana:

Kontekstuālas vērtību definīcijas un spriedze:

Konceptuālie pierādījumi un pieņēmumi:

Empīriski pierādījumi, dalībnieki un ierobežojumi:

Tehniskās iespējas un ietekme uz vērtībām:

[Avoti]

„Ethical Engineer” apstiprināja KA220-HED pieteikumu, WP3 „Mācīšana un mācīšanās”.

https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf. <https://standards.ieee.org/standard/7000-2021.html>

Turpinājums nākamajā slaidā



Obligātās prasības:

Prasības, pieņemšanas testi un atbildīgie:

Lēmums: TURPINĀT PĀRSKATĪT APSTĀJIES

Aizsardzības pasākumi, atlikušās nesaskaņas un domstarpības:

Uzraudzība, korektīvie pasākumi un pārskatīšanas iemesli:

[Avoti]

„Ethical Engineer” apstiprināja KA220-HED pieteikumu, WP3 „Mācīšana un mācīšanās”.

https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf. <https://standards.ieee.org/standard/7000-2021.html>

Lomu spēle: projekta pārskatīšana, ņemot vērā izmaksu un grafika spiedienu

- Eksoskeleta piegādātājs var veikt piegādi laikā tikai tad, ja paliek ieslēgta mākoņanalītika. Vadība apgalvo, ka vēlāk ieviesta politika var novērst ļaunprātīgu izmantošanu. Darbinieku pārstāvji pieprasa tehnisko nošķiršanu pirms jebkāda izmēģinājuma. Drošības komandai ir nepieciešama adaptīvā pielāgošanas funkcija augsta riska uzdevumiem.
- Lomas: inženierijas vadītājs, drošības inženieris, darbinieku pārstāvis, datu vadītājs, piegādātājs un neatkarīgais pārbaudītājs.
- Izlemiet, kuras vērtības ir obligātas, vai politika ir pietiekama, kura opcija ir īstenojama un kurš var aizkavēt izlaišanu. Reģistrējiet vienu nepieņemamu argumentu un vienu pamatotu iemeslu ierobežotam izmēģinājumam.

Individuāla pārdomāšana un nodošana

- Uzrakstiet īsu inženierijas argumentu.
- Kura ieinteresētā puse vai vērtība sākotnēji tika izlaista?
- Kādi pierādījumi mainīja problēmas definīciju vai dizainu?
- Kura prasība padarīja vērtību pārbaudāmu?
- Kāda robežvērtība novērsa nepieņemamu kompromisu?
- Kas uzņemas atlikušo slogu vai risku?
- Kāda izmaiņa liktu atkārtoti izskatīt lēmumu?

Pārnesiet metodi uz citu inženierzinātņu disciplīnu un norādiet vienu izmaiņu ieinteresētajās pusēs, pierādījumos vai pārbaudē.

Tests

1. Kā tiešais ieinteresētais atšķiras no netiešajam ieinteresētajam?
2. Kādi ir trīs vērtībai jutīgā projektēšanas pētījumi?
3. Kāpēc vērtības jādefinē kontekstā?
4. Kas saista vērtību ar inženierijas pieņemšanas testu?
5. Kāpēc pirms svērtu punktu piešķiršanas tiek piemērotas obligātās robežvērtības?
6. Nosauciet vienu risku, kas saistīts ar tehnisku intervenci, piemēram, SMOTE, PCA vai post hoc paskaidrojumu.
7. Kādi pierādījumi liecina, ka ieinteresēto pušu līdzdalība ietekmēja dizainu?

[Avoti]

- https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf
- <https://standards.ieee.org/standard/7000-2021.html>
- Oriģinālā 5. moduļa prezentācija, ko nodrošinājis projekta „Ethical Engineer” konsorcijs.

Atbildes uz viktorīnu un mācību pamatojums

1. Tiešie ieinteresētie dalībnieki mijiedarbojas ar sistēmu vai saņem no tās lēmumus; netiešie ieinteresētie dalībnieki tiek ietekmēti caur plašākām sekām.
2. Konceptuāli, empīriski un tehniski pētījumi.
3. Drošība, autonomija vai taisnīgums var nozīmēt dažādas lietas atkarībā no cilvēkiem, prakses un sistēmām.
4. Izsekojama vērtību hierarhija: vērtība, norma, prasība, pieņemšanas kritērijs un pierādījums.
5. Svērtais koprezultāts nedrīkst atcelt nopietnu juridisku, drošības vai tiesību pārkāpumu.
6. Pamatotas atbildes ietver kalibrēšanas izmaiņas, informācijas zudumu, ievestu neobjektivitāti, nestabilitāti vai maldinošu paskaidrojumu.
7. Reģistrētas izmaiņas problēmā, prasībā, variantā, testā vai lēmumā, kas saistītas ar ieinteresēto pušu pierādījumiem.

Mācību rezultātu pārbaude

Pirms pabeigšanas pārlicinieties, ka varat:

1. Kartējiet tiešās un netiešās ieinteresētās puses, varu un izslēgšanu.
2. veikt konceptuālus, empīriskus un tehniskus pētījumus.
3. pārvērst vērtības prasībās un pieņemšanas kritērijos.
4. salīdzināt datus, modeļus, saskarnes un sistēmu iespējas, izmantojot sliekšņus un pierādījumus.
5. Izstrādājiet uz vērtībām balstītu projektēšanas dokumentāciju, iekļaujot atlikušās nesaskaņas un pārskatu.

Pierādījumi:

Ieinteresēto pušu karte, vērtību hierarhija, variantu salīdzinājums un projektēšanas dokumentācija.

[Avoti]

- „Ethical Engineer” apstiprināja KA220-HED pieteikumu, WP3 „Mācīšana un mācīšanās”.
- https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf
- <https://standards.ieee.org/standard/7000-2021.html>

leguldījums kursa dokumentācijā

Pievienojiet šos 5. moduļa rezultātus savai ētiskās inženierijas dokumentācijai:

1. Ieinteresēto pušu, varas un izslēgšanas karte.
2. Kontekstuālas vērtību definīcijas un vērtību pretrunas.
3. Konceptuāli, empīriski un tehniski pierādījumi.
4. Vērtību hierarhija ar prasībām un pieņemšanas testiem.
5. Pierādījumi par datu, modeļu un skaidrojumu intervenci.
6. Obligātās robežvērtības un variantu salīdzinājums.
7. Uz vērtībām balstīti projektēšanas lēmumi, aizsardzības pasākumi un pārskatīšanas iemesli.

Saistiet šos pierādījumus ar darbu par taisnīgumu, privātumu un atbildību, kas aprakstīts 2.–4. modulī.

[Avoti]

- „Ethical Engineer” apstiprināja KA220-HED pieteikumu, WP3 „Mācīšana un mācīšanās”.
- https://nissenbaum.tech.cornell.edu/papers/Value_Sensitive_Design.pdf
- <https://standards.ieee.org/standard/7000-2021.html>



Paziņojums par finansējumu un izmantošanu izglītības nolūkos

Šis atklātais izglītības resurss atbalsta mācīšanu un diskusijas. Tas nav uzskatāms par juridisku, normatīvu, drošības vai profesionālu inženierijas konsultāciju. Pielietojot tā saturu, ievērojiet spēkā esošos likumus, standartus, iestāžu politiku un nozarei specifiskos norādījumus.

Līdzfinansē Eiropas Savienības programma „Erasmus+“.





Ethical Engineer

www.ethicalengineer.eu

Paldies

Jūs esat pabeidzis 5. moduli



Co-funded by
the European Union