



# DATU ANALĪZES IZAICINĀJUMI UN SCENĀRIJI VALSTS PĀRVALDĒ – KĀ PIENEMT DATOS BALSTĪTUS LĒMUMUS

**Aldis Erglis**  
Senior System Architect



*Data platform*

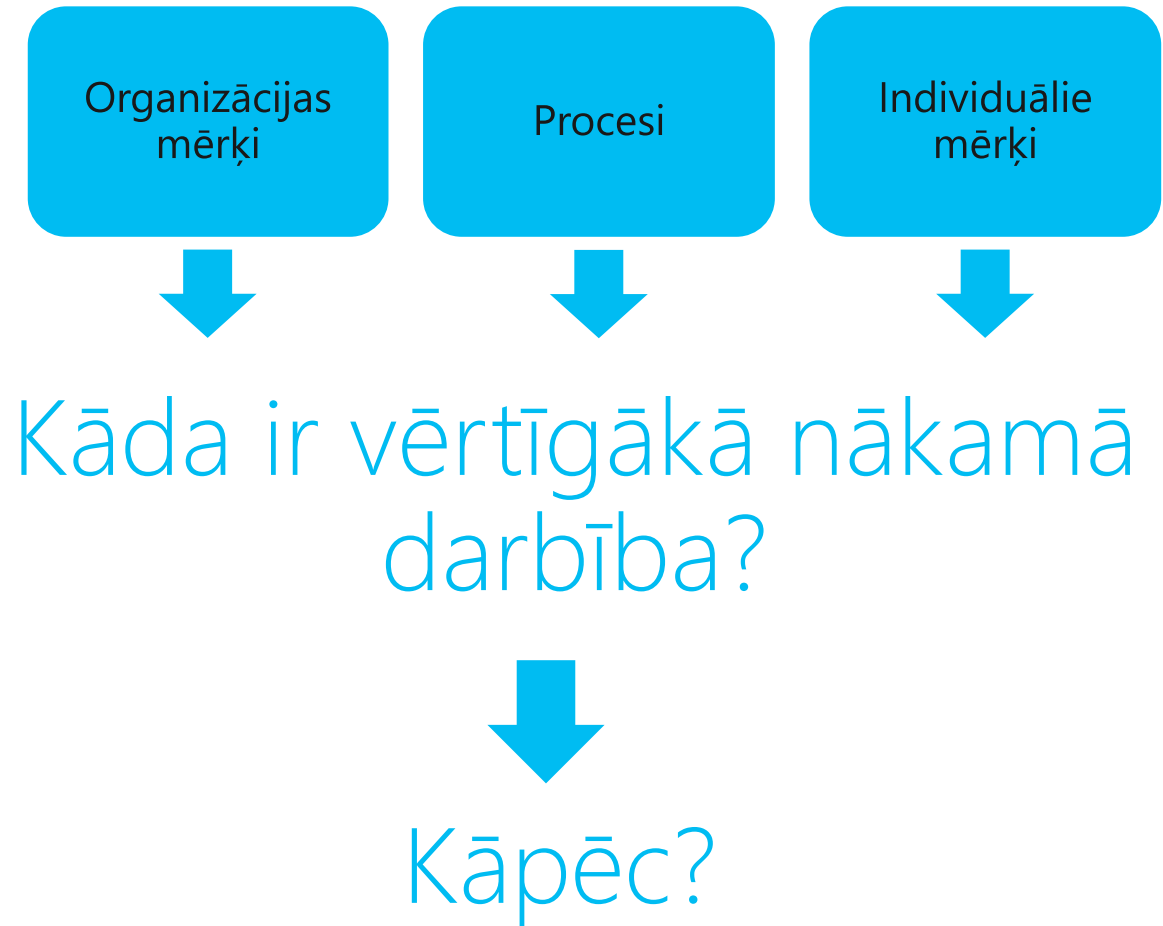
# leguvumi no datu analīzes:

	Pagātne	Tagadne	Nākotne
Informācija	Kas notika? <b>(Atskaites)</b>	Kas notiek tagad? <b>(Brīdinājumi)</b>	Kas notiks? <b>(Ekstrapolācija)</b>
Izpratne	Kāpēc tas notika? <b>(Modelēšana)</b>	Kāda ir nākamā labākā aktivitāte? <b>(Ieteikumi)</b>	Kas ir labākais/ sliktākais, kas var notikt? <b>(Prognozēšana)</b>



Thomas H.  
Davenport

# Rekomendāciju sistēmas



Jau strādā IT tehnoloģiju jomā:

amazon

NETFLIX

Google

Baidu 百度

Alibaba.com

Microsoft

Tiek izstrādāti risinājumi procesu  
Pārvaldības un lēmumu pieņemšanas  
Jomā!

# Priekšnoteikumi rekomendāciju sistēmu attīstībai valsts sektorā

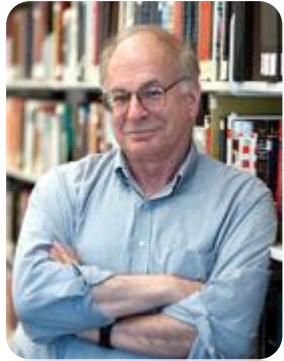
Tehnoloģijas,  
algoritmi

Zināšanas un  
datu analītiķi

Dati un biznesa  
uzdevumi

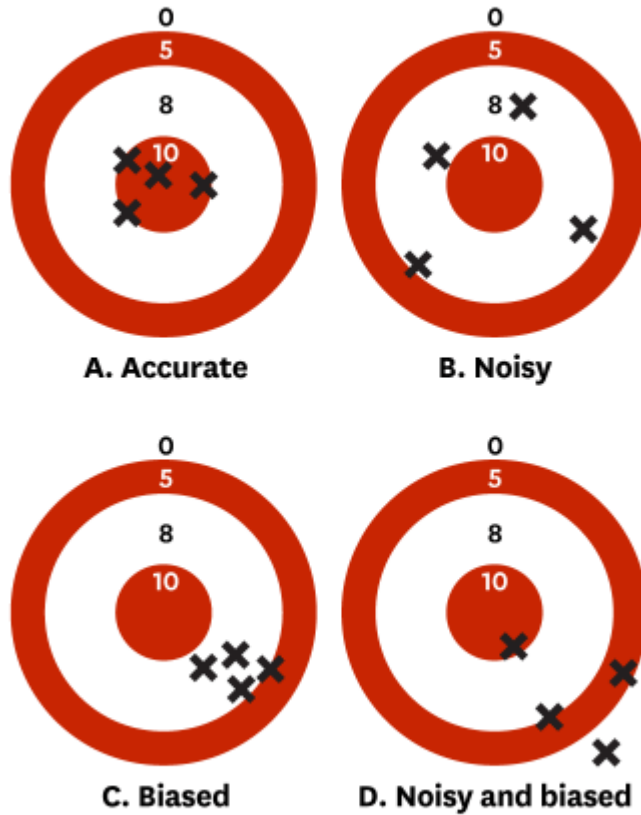
Mākslīgā intelekta stratēģija

# DEEP LEARNING – Cilvēka līmenis



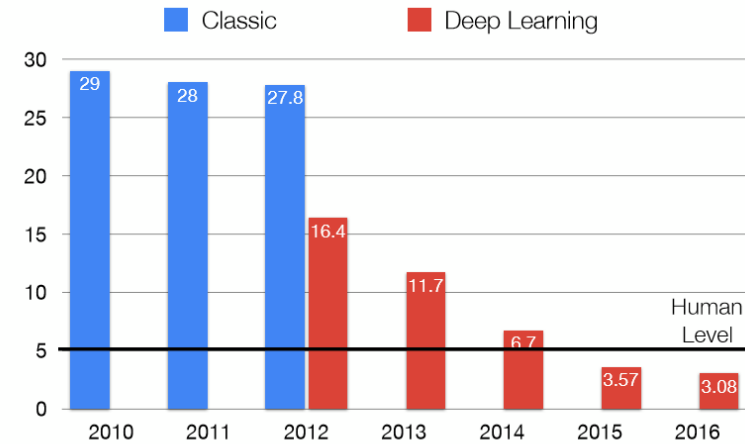
Daniel Kahneman

## How Noise and Bias Affect Accuracy



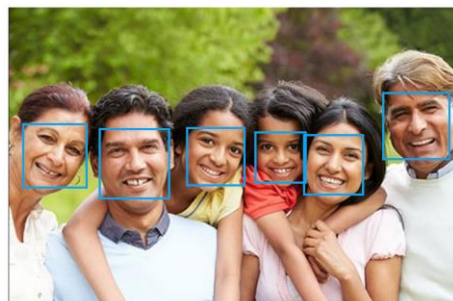
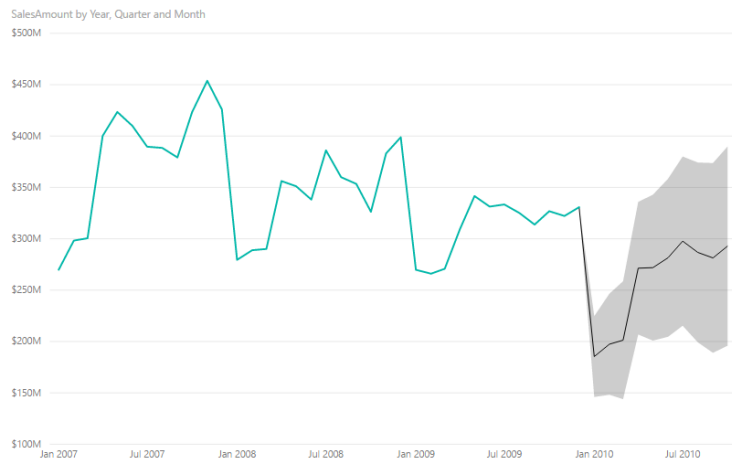
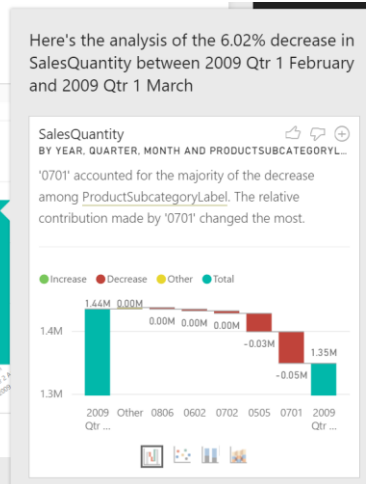
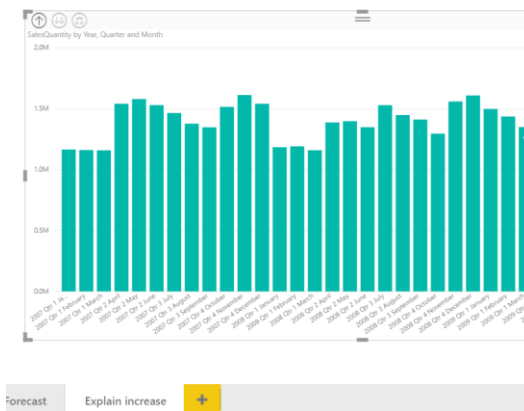
SOURCE DANIEL KAHNEMAN,  
ANDREW M. ROSENFELD,  
LINNEA GANDHI, AND TOM BLASER  
FROM "NOISE," OCTOBER 2016

© HBR.ORG





# Tehnoloģiju un algoritmu demokratizācija



```

Detection Result:
6 faces detected

JSON:
[
  {
    "faceRectangle": {
      "left": 89,
      "top": 128,
      "width": 71,
      "height": 71
    },
    "scores": {
      "anger": 3.44697156e-7,
      "contempt": 6.8442256e-7,
      "disgust": 0.00089313403,
      "fear": 2.98478118e-13,
      "happiness": 0.999748647,
      "neutral": 0.000157187445,
    }
  }
]
    
```

## Vision

- Computer Vision
- Content Moderator
- Emotion
- Face
- Video

## Speech

- Bing Speech
- Custom Recognition
- Speaker Recognition

## Language

- Bing Spell Check
- Language Understanding
- Linguistic Analysis
- Text Analytics
- Translator
- WebLM

## Knowledge

- Academic
- Entity Linking
- Knowledge Exploration
- QnA Maker
- Recommendations

## Search

- Bing Autosuggest
- Bing Image Search
- Bing News Search
- Bing Video Search
- Bing Web Search

# Expert Systems

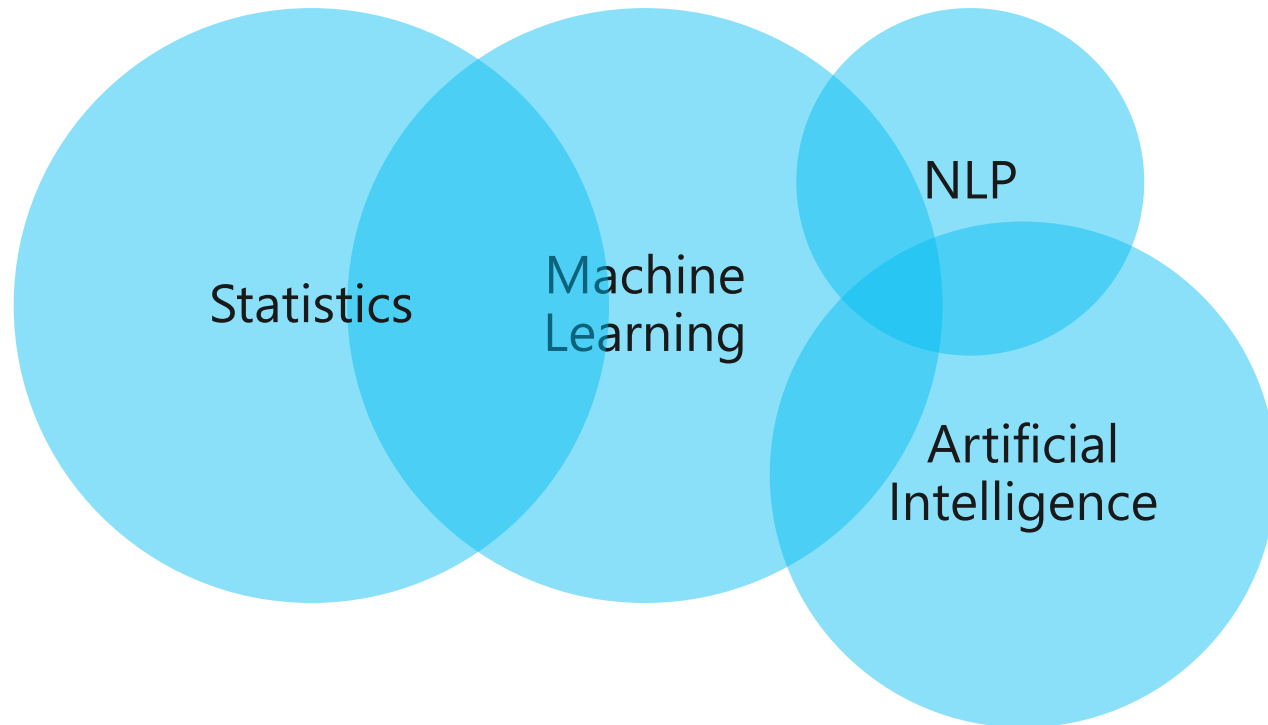
- Zināšanu bāze (noteikumi, IF-THEN)
- Faktu datu bāze (Konteksts)
- Aprēķina mehānisms (Secinājumi)
- Izskaidrošanas mehānisms



# Kas ir Machine Learning

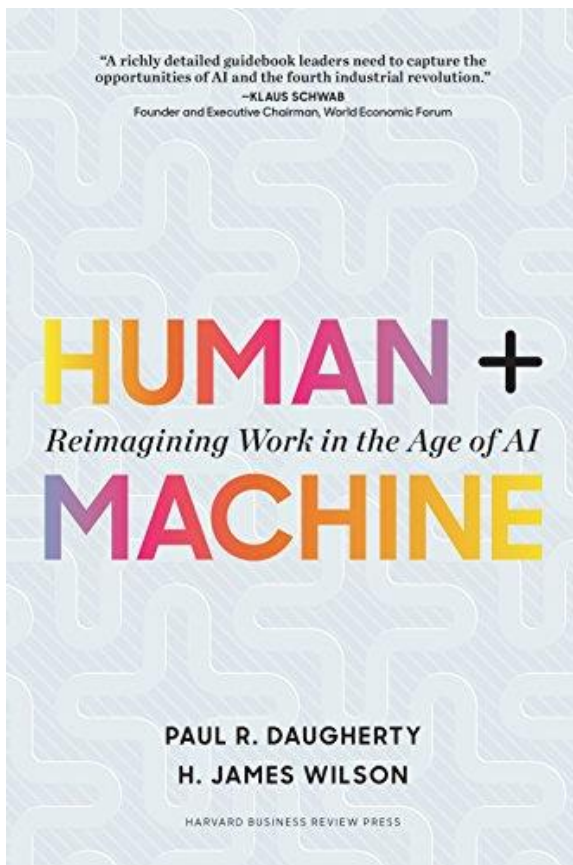
Machine learning is the subfield of [computer science](#) that, according to [Arthur Samuel](#) in 1959, gives "computers the ability to learn without being explicitly programmed."

- Descriptive
- Prescriptive
- Predictive





# Vai algoritmi aizvietos darbiniekus?



## The missing middle—left side

Lead	Empathize	Create	Judge	Train	Explain	Sustain	Amplify	Interact	Embody	Transact	Iterate	Predict	Adapt
H Human-only activity				Humans complement machines			AI gives humans superpowers			M Machine-only activity			
				Human and machine hybrid activities									


# 1 miljona vērts jautājums – atrisināts hakatonā

## Netflix Prize


2009. Gads 21. septembrī 1 miljona dolāru balvu saņem BellKor Pragmatic Chaos komanda, kas uzlabo Netflix filmu ieteikšanas algoritmu par 10.06%.



# Zināšanas un datu analītiķi

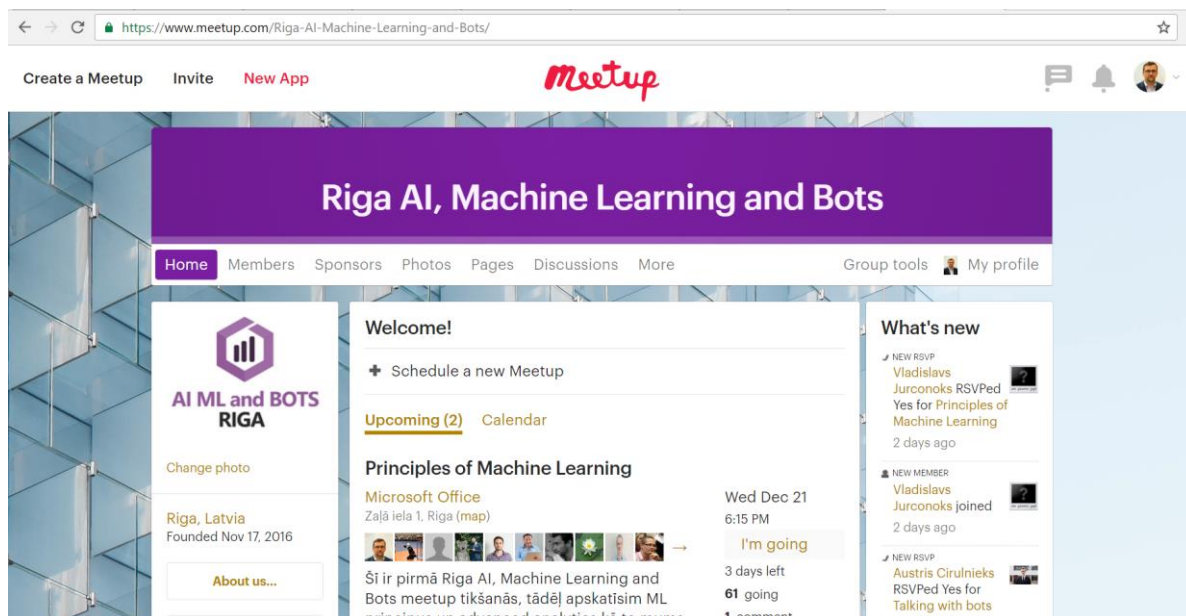


LATVIJAS  
UNIVERSITĀTE  
ANNO 1919



Microsoft

Datu analītikas izaicinājums:  
izaicinājumi publiskajā pārvaldē




https://www.meetup.com/Riga-AI-Machine-Learning-and-Bots/

Create a Meetup Invite New App

meetup

## Riga AI, Machine Learning and Bots

Home Members Sponsors Photos Pages Discussions More Group tools My profile



**AI ML and BOTS  
RIGA**

Change photo

Riga, Latvia  
Founded Nov 17, 2016

About us...

Welcome!

+ Schedule a new Meetup

Upcoming (2) Calendar

**Principles of Machine Learning**

Microsoft Office  
Zaļā iela 1, Rīga (map)

Wed Dec 21  
6:15 PM

I'm going

3 days left

61 going

1 comment

What's new

NEW RSVP  
Vladislavs Jurconoks RSVPed Yes for Principles of Machine Learning 2 days ago

NEW MEMBER  
Vladislavs Jurconoks joined 2 days ago

NEW RSVP  
Austrijs Cirulnieks RSVPed Yes for Talking with bots



[www.meetup.com](http://www.meetup.com)



# Principles of Machine Learning

2016. gada 21. Decembris 18:15

Rīgā, Zaļā iela 1  
Microsoft Latvija biroja



# Talking with Bots

2017. gada 11. Janvāris 18:15

Rīgā, Zaļā iela 1  
Microsoft Latvija biroja



# Deep learning Reinforcement learning

2017. gada 16. Februāris 18:15

Rīgā, Zaļā iela 1  
Microsoft Latvija biroja



# eStepControl ML praktisks pielietojums

2017. gada 9. Marts 18:15

Rīgā, Zaļā iela 1  
Microsoft Latvija biroja



# Deep learning risinājumi gudrākām pilsētām

2017. gada 6. Aprīlis 18:15

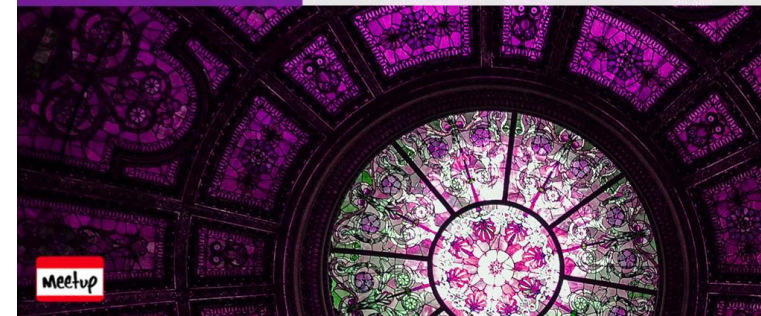
Rīgā, Zaļā iela 1  
Microsoft Latvija biroja



# Deep Learning - tehnikais ieskats, kā tas darbojas

2017. gada 25. Maijs 18:15

Rīgā, Zaļā iela 1  
Microsoft Latvija biroja



# Computer Vision & more

2017. gada 27. Aprīlis 18:15

Rīgā, Zaļā iela 1  
Microsoft Latvija biroja



# How to setup advanced analytics process

2017. gada 16. Jūnijs 18:15

Rīgā, Zaļā iela 1  
Microsoft Latvija biroja



# Generative modeling with Convolutional Neural Networks

2017. gada 3. Augusts 18:15

Rīgā, Zaļā iela 1  
Microsoft Latvija biroja





# Kā saprast ML:



Tiesu administrācija

Datu kopā atrodas dati par 2017. gadā pirmās instances tiesās izskatītajām 3. kategorijas civillietām (Prasības par zaudējumu un parāda piedziņu).

	A	B	C	D	E	F	G	H	I	J	
1	Lietas ID (caseID)	Tiesas ID (courtID)	Dzimšanas datums	Ienākšanas kārtība	Ienākšanas datums	Izskatīšanas datums	Būtība	Prasījuma apmērs	atb_fiz	atb_jur	pr
2	213477	40	060159	Pirmo reizi	6/5/2003	1/23/2017	parāda piedziņa	0.00	0	1	
3	448971	32	050767	Pirmo reizi	11/29/2005	4/6/2017	zaudējumu piedziņa	0.00	1	0	
4	477470	279	010551	Sakarā ar tiesu apvienošanu	1/31/2017	10/2/2017	parāda piedziņa	21414.34	2	0	
5	575613	27	030463	Pirmo reizi	6/4/2007	10/26/2017	zaudējumu piedziņa	916.25	6	0	
6	580131	32	250174	Pēc piekritības no citas tiesas	10/7/2015	11/13/2017	parāda piedziņa	1588192.95	4	0	
7	614723	279	110274	Sakarā ar tiesu apvienošanu	1/31/2017	6/19/2017	parāda piedziņa	0.00	0	1	
8	619189	25	010551	Pirmo reizi	10/26/2007	1/3/2017	zaudējumu piedziņa	0.00	0	2	
9	623711	25	181271	Pirmo reizi	11/9/2007	1/12/2017	parāda piedziņa	0.00	0	1	
10	628048	30	071058	Pirmo reizi	11/26/2007	11/16/2017	par rokas naudas piedziņu	20.00	0	1	
11	686457	279	260765	Sakarā ar tiesu apvienošanu	1/31/2017	2/21/2017	parāda piedziņa	0.00	0	1	
12	687481	26	050856	Pirmo reizi	5/16/2008	3/29/2017	zaudējumu piedziņa	177.21	0	1	
13	707297	4	221265	Pirmo reizi	7/3/2008	3/2/2017	parāda piedziņa				
14	753938	279	110274	Sakarā ar tiesu apvienošanu	1/31/2017	12/18/2017	parāda piedziņa				
15	776387	26	290568	Pirmo reizi	11/5/2008	3/14/2017	zaudējumu piedziņa				
16	781658	4	221265	Pirmo reizi	11/14/2008	3/2/2017	parāda piedziņa	2269.75	0	1	
17	781661	4	221265	Pirmo reizi	11/14/2008	3/2/2017	parāda piedziņa	479.74	0	1	

Virs 10000 ierakstu



# Paredzamais rezultāts

Šo datu apgrozības rezultāts ir spēt prognozēt lietas izskatīšanas ilgumu.

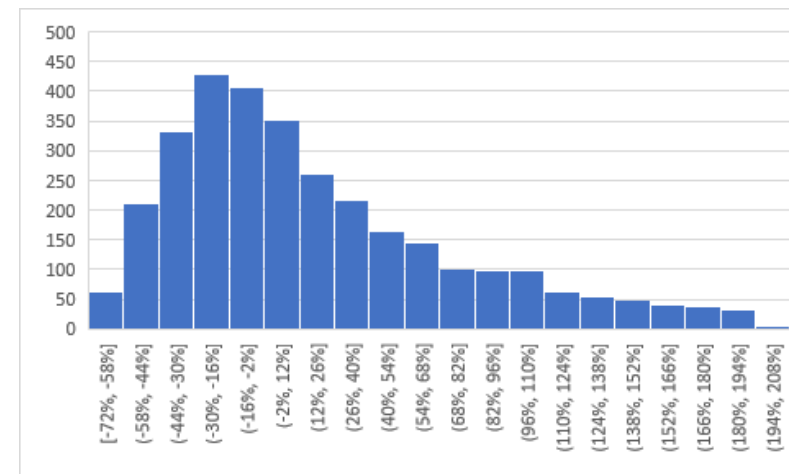
- Tiesas ID
- Lietu skatījušā tiesneša vecums
- Kad pirmo reizi lieta parādās tiesu sistēmā
- Lietas saņemšanas datums
- Lietas pabeigšanas datums
- Par ko bijusi lieta
- Prasījuma apmērs lietā
- Atbildētāju un prasītāju personu skaits lietā
- Cik sēdes bijušas nozīmētas
- Cik no sēdēm notikušas
- Vai kāds no lietas dalībniekiem ir sodīts ar naudas sodu
- Vai atbildētājs dzīvo ārzemēs
- Vai lietas dalībnieks ir atbrīvots no nodevas



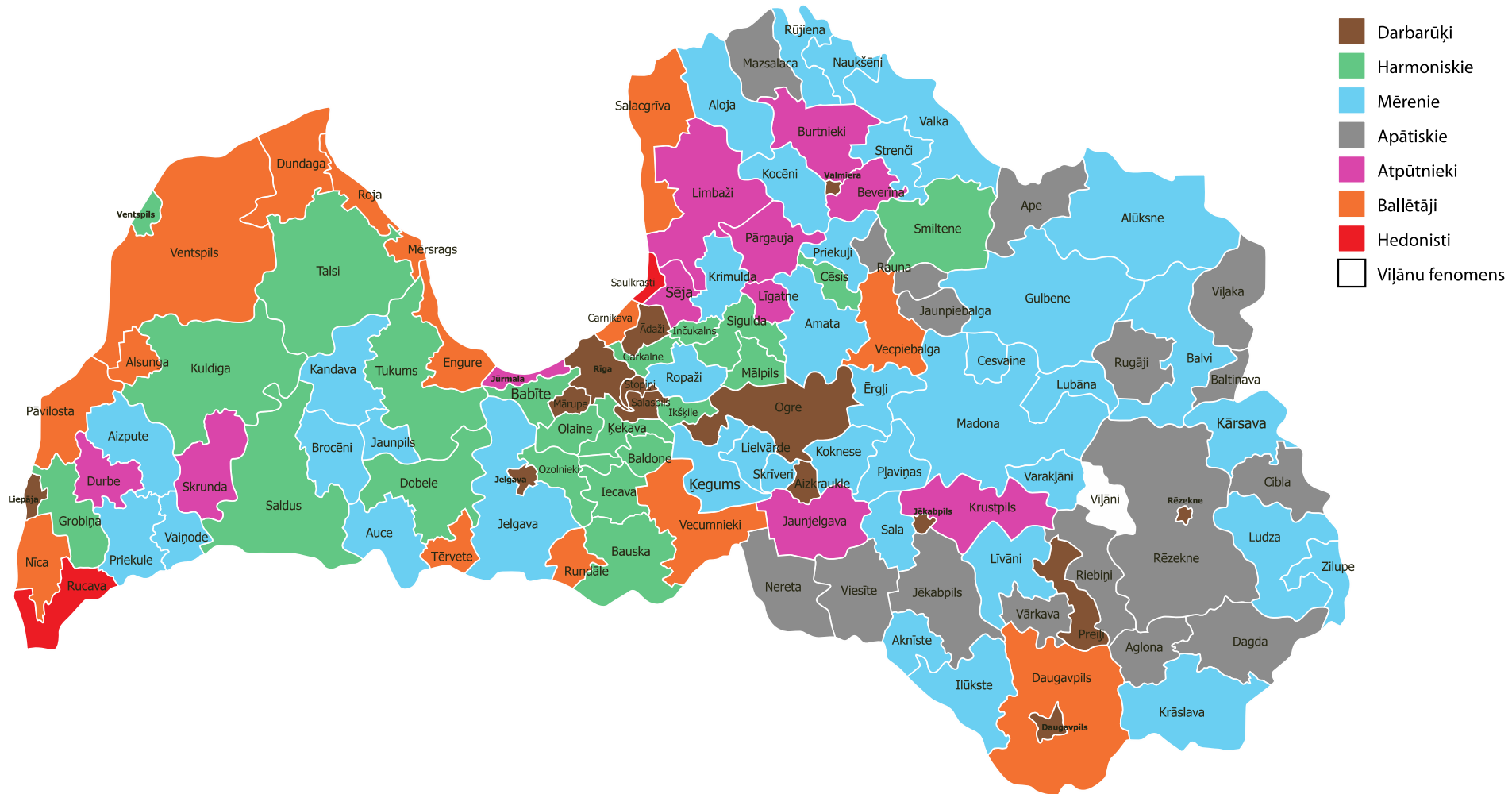
**Mēnešu skaits**  
(izskatīšanas  
ilgums)

# Rezultāti

Lietas ID	Īstais ilgums (Mēneši)	Prognozētais ilgums (Mēneši)	Delta	Delta(%)
1916743	7.8	4.0	-3.8	-49%
1878423	10.6	6.2	-4.3	-41%
1840955	15.3	15.5	0.2	1%
1977437	1.9	4.1	2.1	110%
1963272	3.5	5.4	1.9	54%
1912862	6.1	8.1	1.9	32%
1900793	7.9	6.0	-1.9	-24%
2063007	2.5	4.1	1.6	65%
1987265	5.7	8.6	2.9	51%
1940923	4.7	4.8	0.1	2%
1971195	5.8	4.3	-1.5	-26%
1956663	2.9	5.7	2.9	100%
1914134	15.5	9.9	-5.6	-36%
1967845	6.9	4.9	-2.0	-29%
2060527	2.1	4.1	2.0	95%
1971385	4.2	4.8	0.6	14%
1969387	7.7	3.8	-3.9	-50%

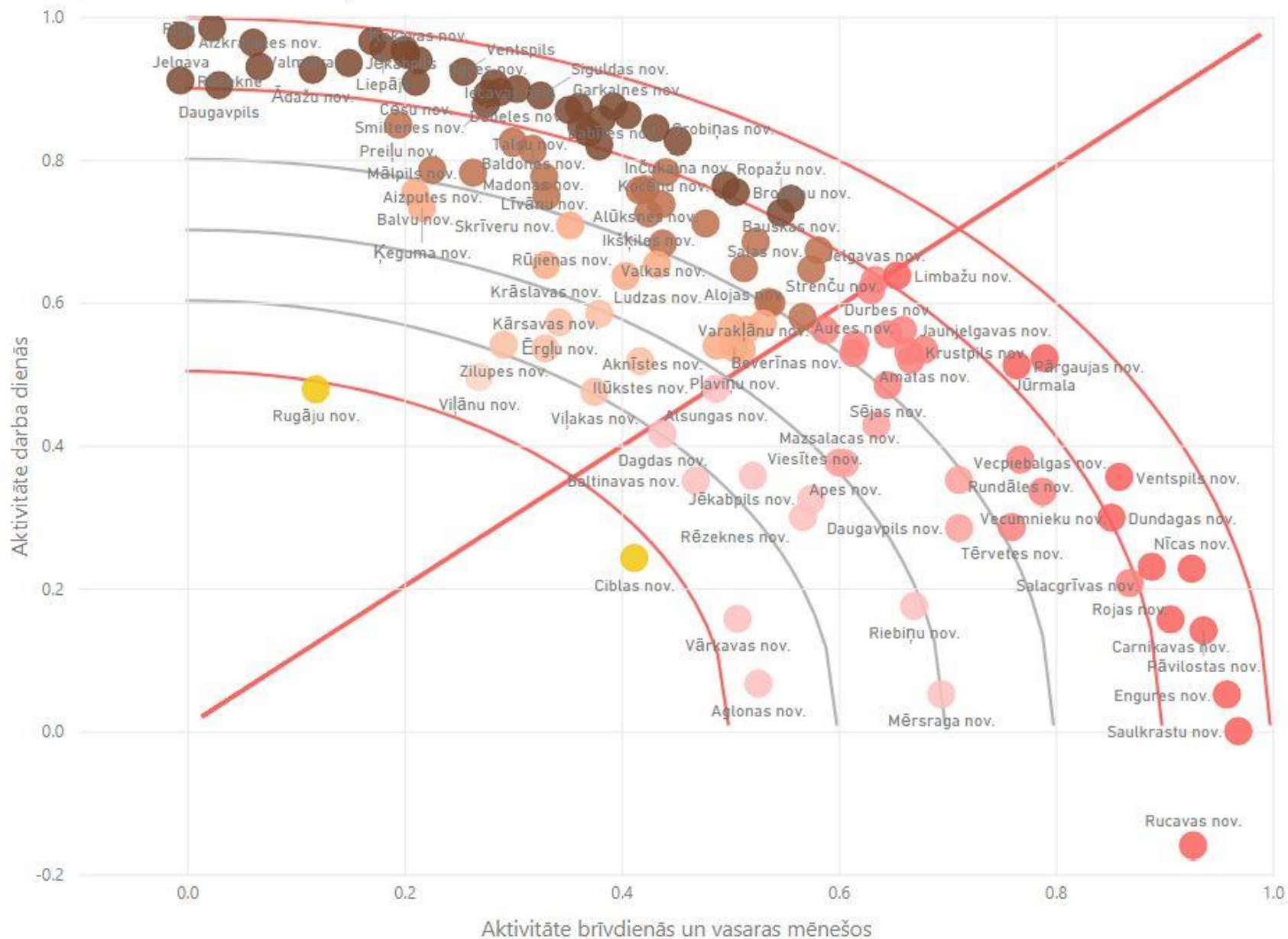


# Reģionu ekonomiskās attīstības indekss



# Efektivitāte 2017. gads

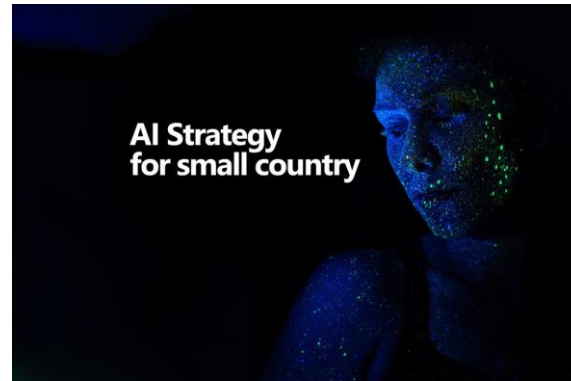
Brīvdienas, Darba dienas and Efektivitāte by Novads and ID



# Priekšnoteikumi uz datiem balstītiem lēmumiem



Vai jūsu organizācija ir gatava pieņemt lēmumus balstoties uz datiem?



Mazām valstīm nepieciešama sava mākslīgā intelekta stratēģija

- Pareizu datu uzskrāšana
- Vadītāji ar datu analīzes kompetencēm
- Latviešu valoda
- Mākslīgā intelekta stratēģija

<https://www.linkedin.com/in/aldiserglis/>

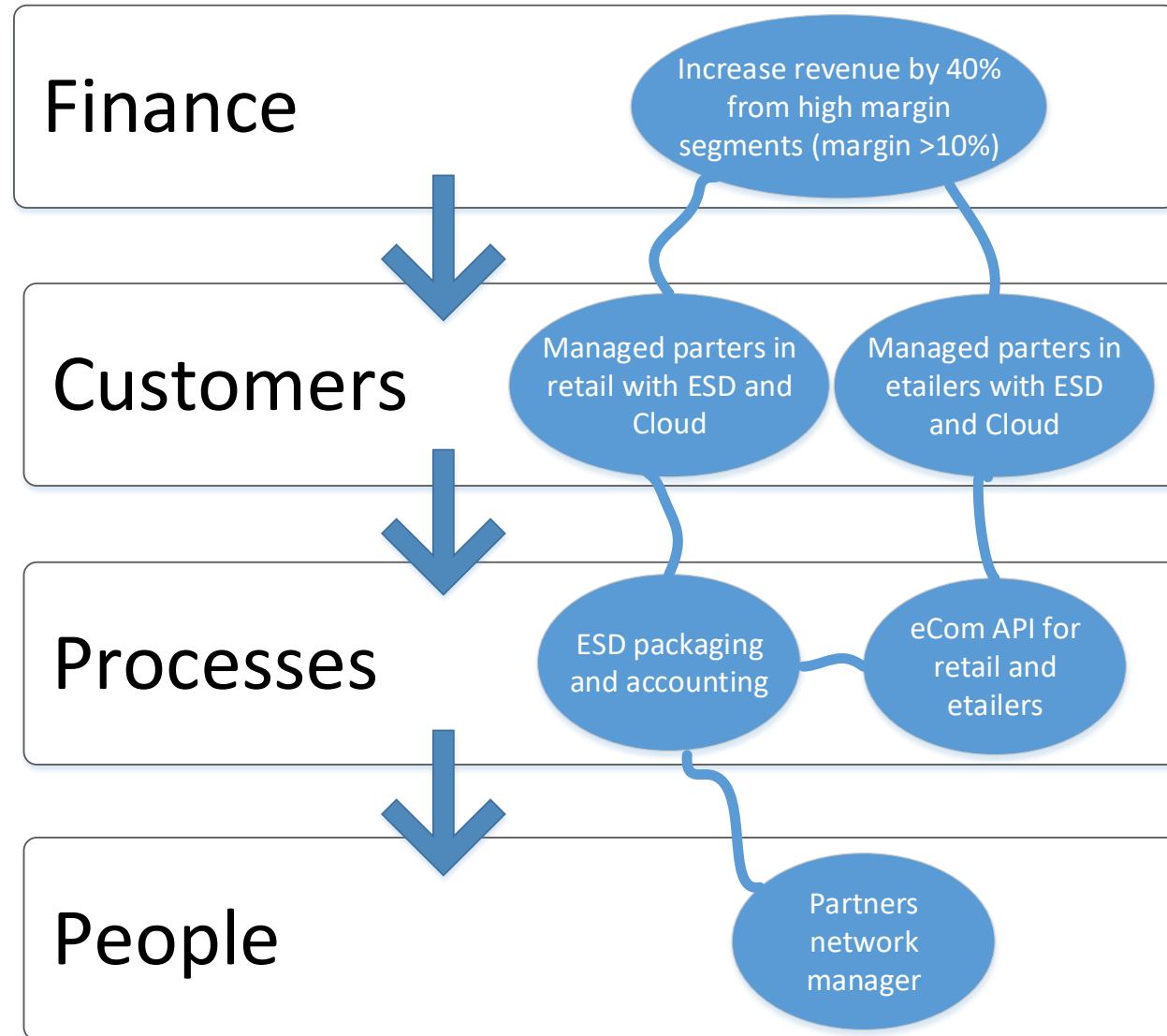


Pārāk daudz iespēju – alternatīvās izmaksas

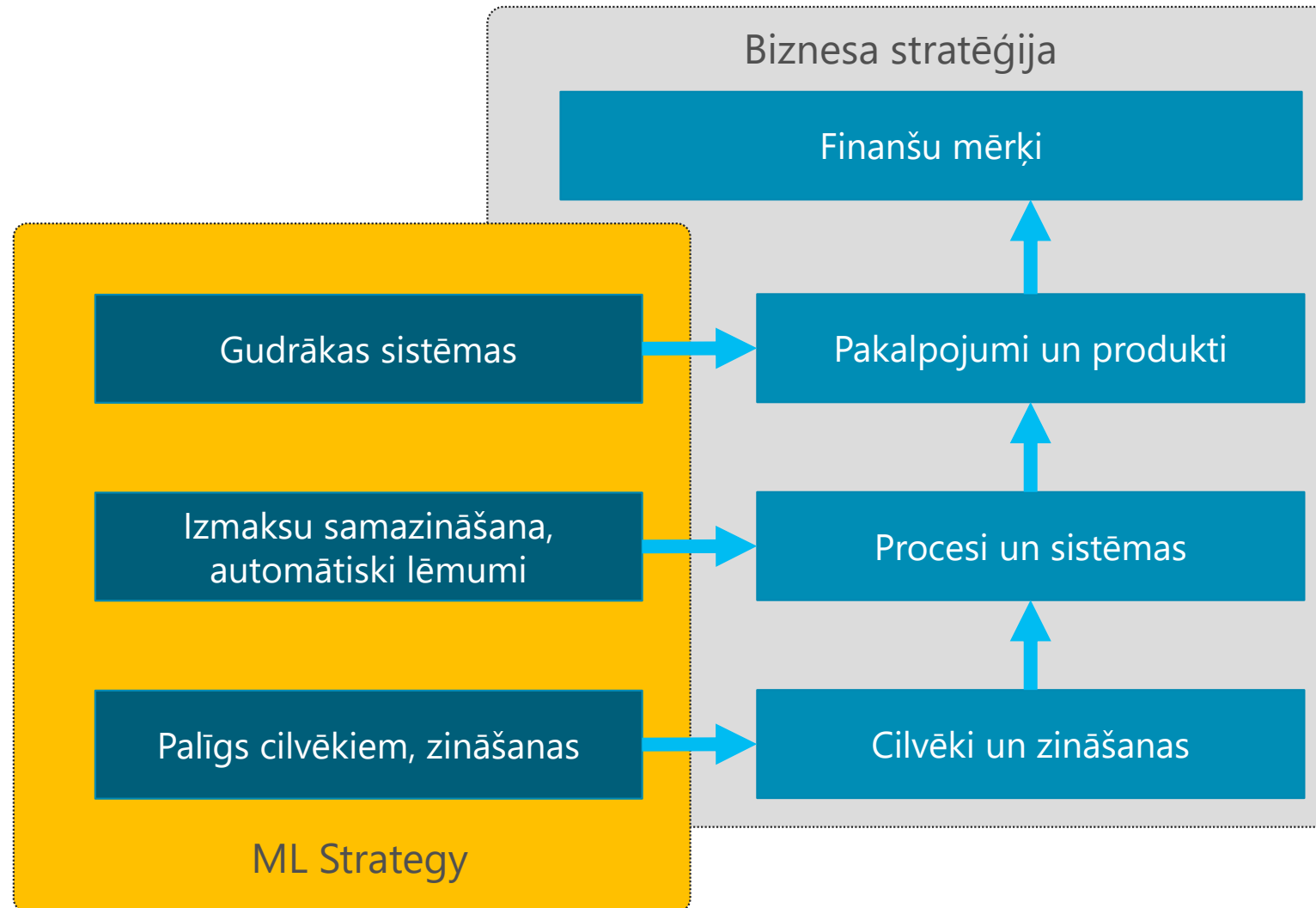
# Nepieciešama organizācijas ML Stratēģija

Universālas,  
eksperimentālas ML  
implementācijas  
negarantē stratēģiskus  
ieguvumus

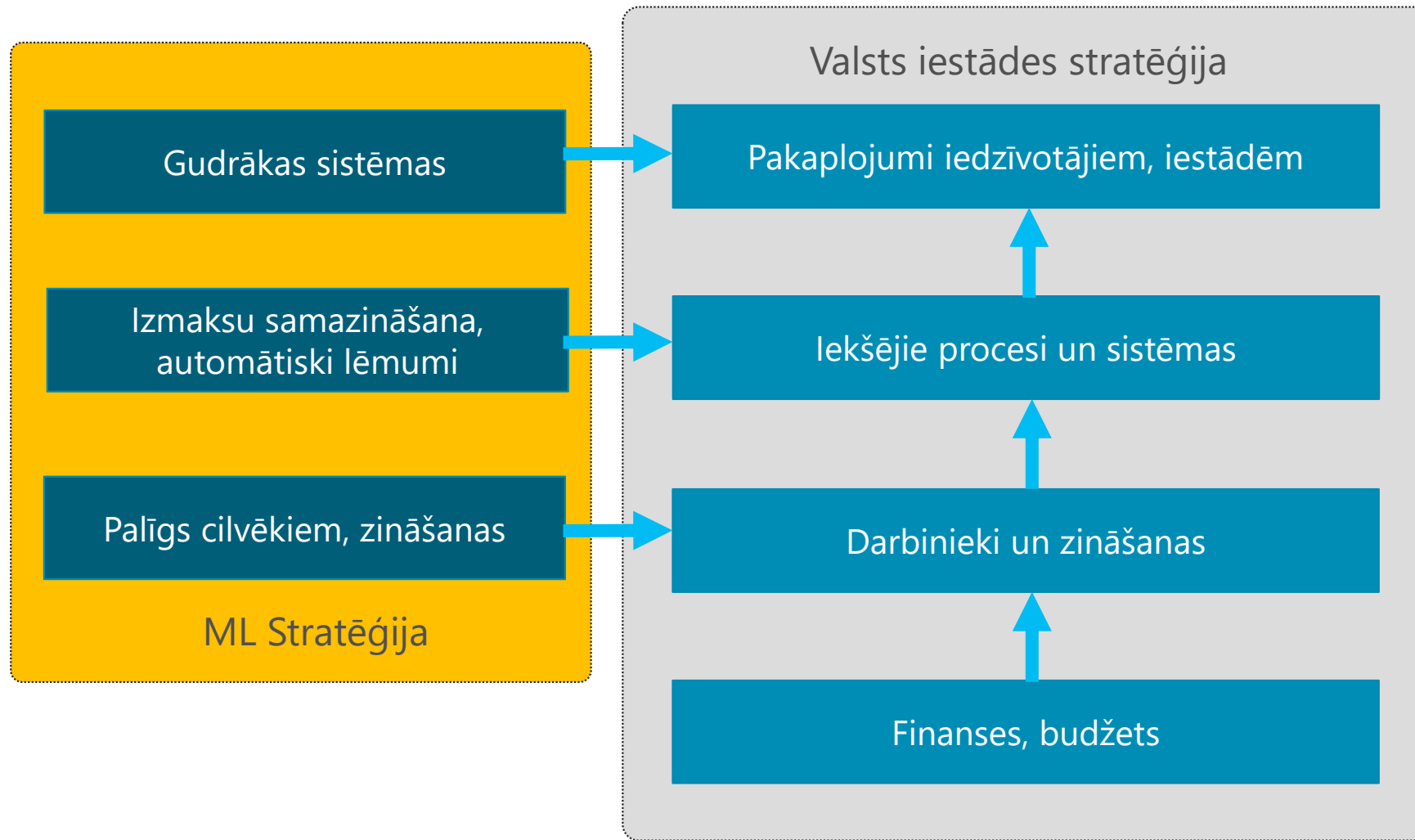
# Biznesa stratēģijas dekompozīcija



# AI/ML stratēģija – atbalsta stratēģija



# Līdzsvarotā mērķu stratēģija valsts sektoram



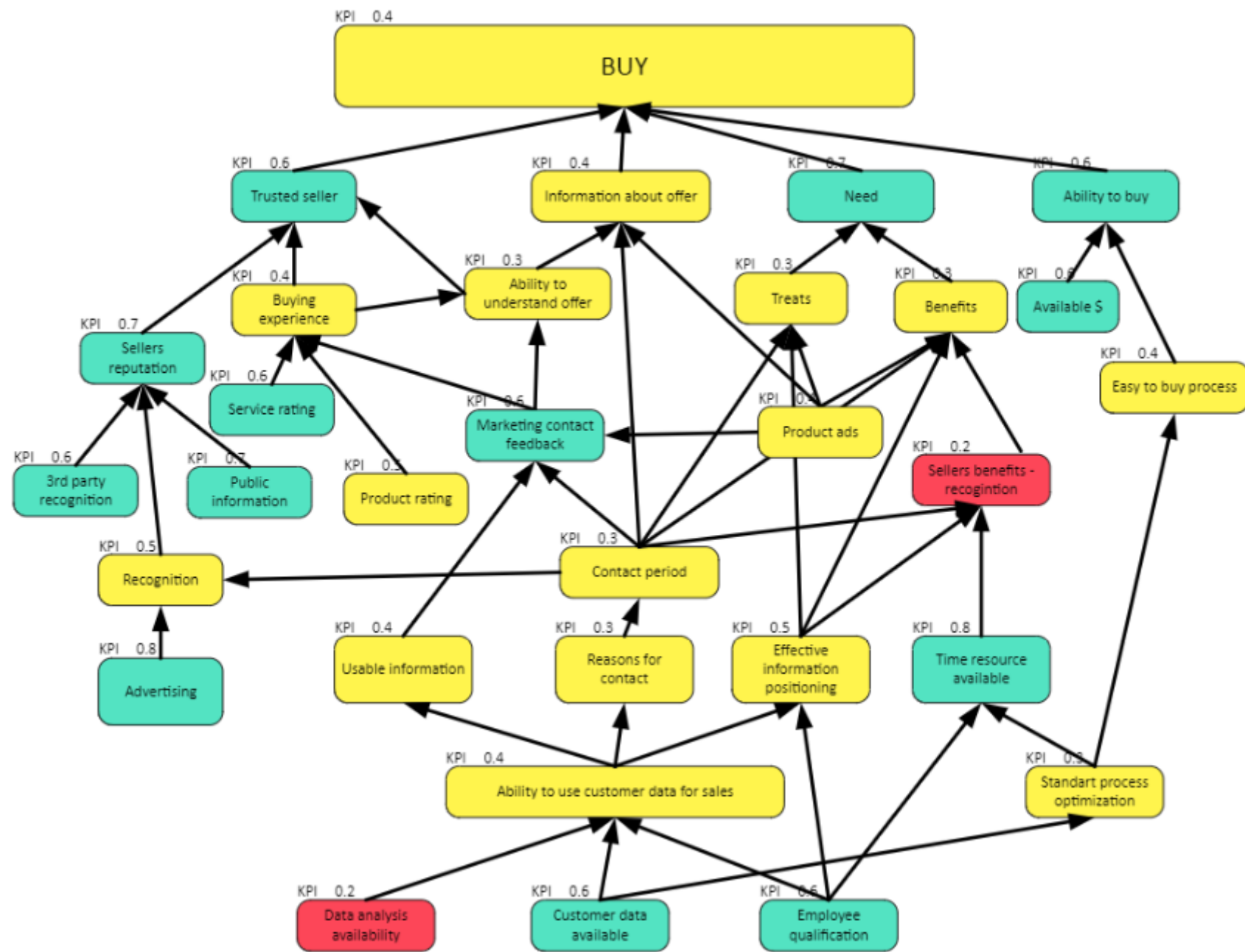
# Datu analīze neatbild uz jautājumu – Kāpēc?

Datos var atrast atbildi uz jautājumiem:

- Cik daudz?
- Kāds, kad?
- Cik bieži?

Pat lielie dati nedod atbildi uz jautājumu – Kāpēc?





Country

- Estonia
- Latvia

# Cēloņsakarību izgūšana no procesiem

- Katrā organizācijā ir vērtību ķēde un procesi

Vērtību ķēde



Biznesa procesi

