



Neuroninių kalbos technologijų taikymai medicinos bei teisės srityse

dr. Darius Amilevičius
Vytauto Didžiojo universitetas

darius.amilevicius@vdu.lt



Kalbos technologijos VDU

2015 – Semantika1

2020 – Semantika2

2021 – šnekos sintezės ir atpažinimo vystymai

Automatinė teksto analizė

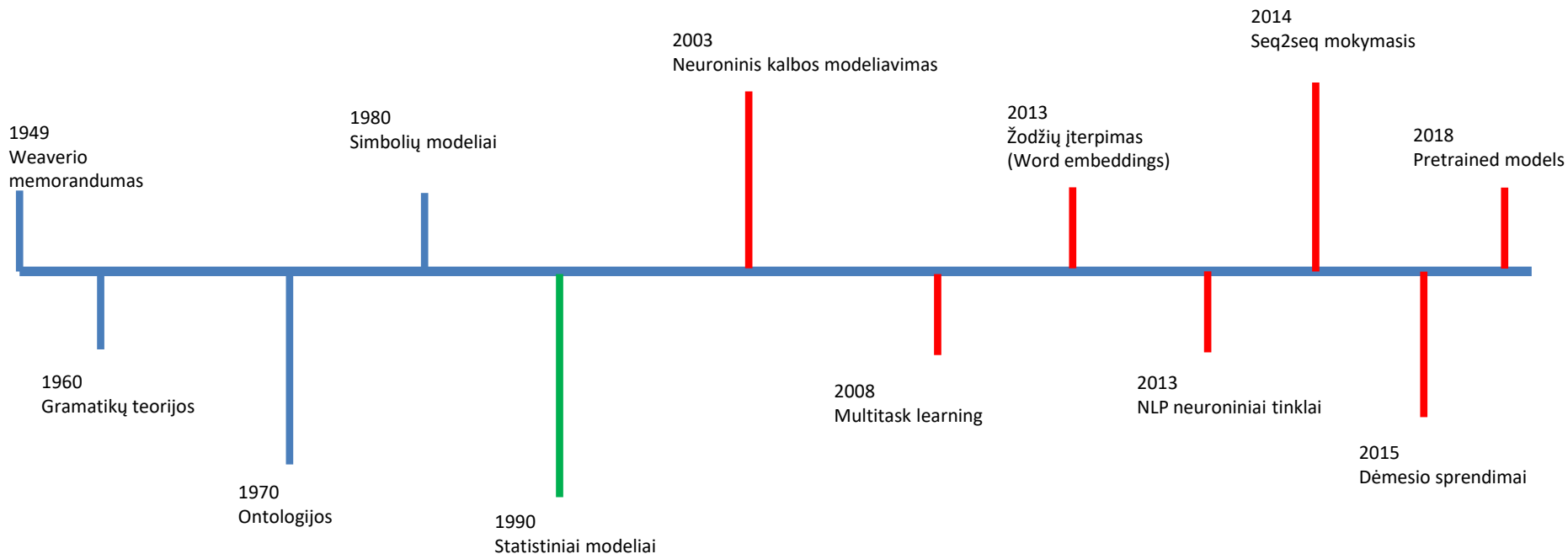
Neuroninis kalbos modeliavimas

.....





Mašinų mokyme skaityti, rašyti, klaustyti ir kalbėti





DI/GM/NLP/NLU

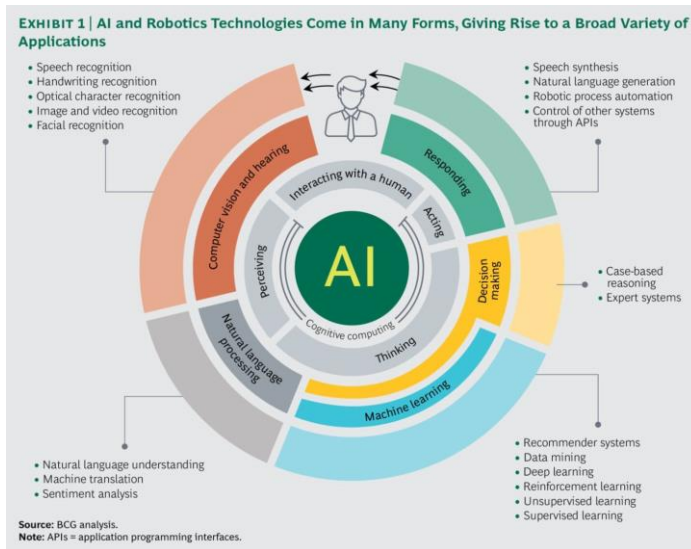
Dirbtinis intelekts (DI)

Natūralios kalbos technologijos (NLP)

EHR
Teisiniai
dokumentai

Nestruktūruoti
duomenys

Intelektualūs
asistentai





Neuroninis kalbos modeliavimas



Žodis = ???

BoW, TF-IDF

Word2vec, Fasttext, Glove, ELMO

```
array([-0.05419922,  0.01708984, -0.00527954,  0.33203125, -0.25
-0.01397705, -0.15039062, -0.265625,  0.01647949,  0.3828125,
-0.03295898, -0.09716797, -0.16308594, -0.04443359,  0.00946045,
 0.18457031,  0.03637695,  0.16601562,  0.36328125, -0.25585938,
  0.375,  0.171875,  0.21386719, -0.19921875,  0.13085938,
-0.07275391, -0.02819824,  0.11621094,  0.15332031,  0.09082031,
 0.06787109, -0.0300293, -0.16894531, -0.20800781, -0.03710938,
-0.22753906,  0.26367188,  0.012146,  0.18359375,  0.31054688,
-0.10791016, -0.19140625,  0.21582031,  0.13183594, -0.03515625,
 0.18554688, -0.30859375,  0.04785156, -0.10986328,  0.14355469,
-0.43554688, -0.0378418,  0.10839844,  0.140625, -0.10595703,
 0.26171875, -0.17089844,  0.39453125,  0.12597656, -0.27734375,
-0.28125,  0.14746094, -0.20996094,  0.02355957,  0.18457031,
 0.00445557, -0.27929688, -0.03637695, -0.29296875,  0.19628906,
 0.20703125,  0.2890625, -0.20507812,  0.06787109, -0.43164062,
-0.10986328, -0.2578125, -0.02331543,  0.11328125,  0.23144531,
-0.04418945,  0.10839844, -0.2890625, -0.09521484, -0.10351562,
-0.0324707,  0.07763672, -0.13378906,  0.22949219,  0.06298828,
 0.08349609,  0.02929688, -0.11474609,  0.00534058, -0.12988281,
```



BERT & Co

“Vandenį iš puodelio į butelį pyliau, kol jis prisipildė.”

“Vandenį iš puodelio į butelį pyliau, kol jis ištuštėjo.”

Kiti

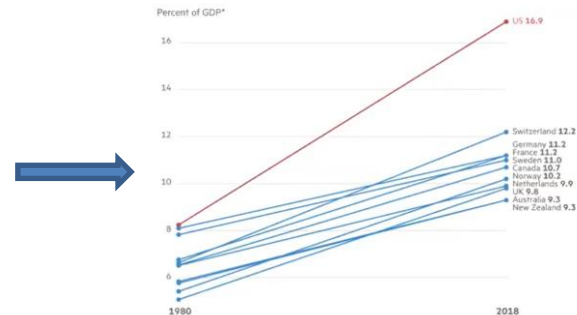


Kontekstas

DI
gamyboje ir pramonėje →



DI
Sveikatos apsaugos ir
teisinėje sistemose →





Kontekstas

- Pasaulyje iki 2050 m. 1 iš 6 žmonių bus vyresni nei 65 metų. Europoje ir Šiaurės Amerikoje šis skaičius bus 1 iš 4.
- PSO – prognozuojama, kad iki 2030 m. bus sukurta 40 milijonų naujų darbo vietų sveikatos apsaugos sektoriuje, bet per šį laikotarpį prognozuojamas 9,9 milijono gydytojų, slaugytojų ir akušerių trūkumas



Kontekstas

2021–2030 METŲ PLĖTROS PROGRAMOS VALDYTOJOS LIETUVOS RESPUBLIKOS TEISINGUMO MINISTERIJOS TEISINGUMOS SISTEMOS PLĖTROS PROGRAMA

Didelis teismų gaunamų bylų skaičius:

1.4.1. Teismai nagrinėja nesudėtingas bylas, kuriose nėra teisinio ginčo

Teismų infrastruktūra neužtikrina teisingumo vykdymo proceso poreikių:

1.5.1. Teismų turimos informacinės technologijos nepadedą užtikrinti reikiamų sąlygų nuotoliniam darbui, teismų veiklos skaitmenizavimui ir informacinio saugumo stiprinimui

Ribotas paslaugų prieinamumas:

3.1.1.1. Paslaugų teikimas tiesioginio kontakto būdu nėra patogus.

3.1.1.2. Rankinio darbo metodai neproporcingai didina procedūras administruojančių specialistų darbo krūvį



Automatizavimas vs. inovacijos vs. sisteminės inovacijos vs. skaitmeninės transformacijos

Remiantis R. Susskindo nuomone, visų pirma būtina atsakyti į penkis esminius klausimus:

- ar technologijos jau pakankamai brandžios, kad galėtų pakeisti gydytoją/teisininką
- ar moraliai jau priimtina, kad mašinos perimtų bent kai kurias gydytojo/teisininko funkcijas
- ar tokios sistemos jau komerciškai perspektyvios, ar jų ekonominė nauda bus didesnė už jų kūrimo išlaidas
- ar tai bus tvaru (nepaskatins atmetimo reakcijos), atsižvelgiant į kultūrinį kontekstą
- ar yra teisinis pagrindas kurti tokias sistemas, dėl ko sprendimo priėmimas iš dalies ar visiškai negali būti patikėtas mašinai

Automatizavimas: sveikatos apsauga

Kūno odoje, daugiau juosmens srityje, numuliarinės plokščios papulės, uždegiminės, su žvynu. Alkūnių, kelių srityje jos susilieja į stambesnius židinius. Tarpgliuteralinėje raukšlėje paraudęs, infiltruotas židiny su įtrūkimu ir maceracija. Kojų nagai distrofiški, padų ir delnų srityje sausos pustulės, padengtos žvynu. Aušpico triada teigiama. Galvoje pavienės plokštelės. Bendras kraujo tyrimas 2003-08-13: leukocitai - $5,5 \times 10^9/l$, eritrocitai - $4,38 \times 10^{12}/l$, hemoglobinas- 134 g/l, trombocitai - $232 \times 10^9/l$, leukocitų formulė: lazdeliniai- 1%, segmentuoti- 55%, bazofilai- 0, eozinofilai- 1%, limfocitai- 33%, monocitai- 10%, ENG pagal Westergreną- 8 mm/val. Šlapimo tyrimas normos ribise. Tyrimas dėl grybelio iš pakitusių rankų nagų: grybelio nerasta. Biocheminis kraujo tyrimas: bilirubinas bendras- 6,2 mkmol/l, ASAT- 23 u/l, ALAT- 21 u/l, gliukozės kiekis kraujyje- 5,7 mmol/l. PUVA terapija, vartojant Oksaraleno po 10 mg 4 tbl 1,5 val prieš PUVA terapijos procedūrą, viso gavo 8 procedūras. Tepėsi deguto 2 % tepalu pusiau su Sinaflanu vieną kartą dnoje



Unstructured data



Text mining



Interoperable & Aggregated



Structured information



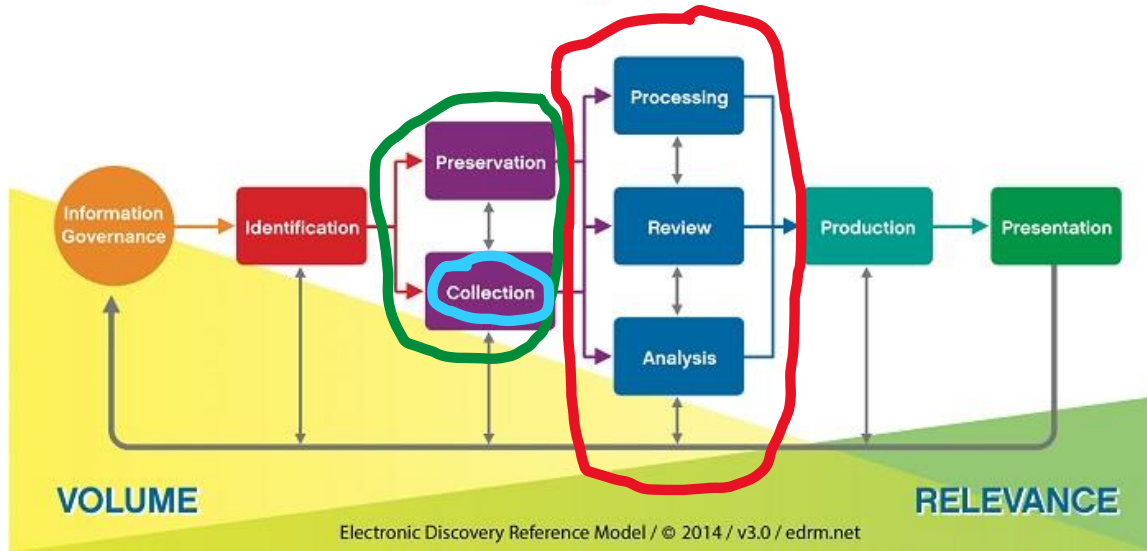
Insights Generation



Actionable knowledge

Automatizavimas: teisinė sistema

Electronic Discovery Reference Model





Problemos

Duomenys

Infrastruktūra



Děkoju už děmesj

