

Prevajalske tehnologije v dobi umetne inteligence

Seminar za Društvo prevajalcev in tolmačev Slovenije

30. 6. 2020

prof. dr. Špela Vintar

Oddelek za prevajalstvo

Filozofska fakulteta, Univerza v Ljubljani

<http://www.lojze.si/spela>

spela.vintar@ff.uni-lj.si

Okvirni program

- Trendi v jezikovni industriji in vpliv Covid-19
- Prevajalska orodja v oblaku
- Popravljanje strojnih prevodov
- Jezikovne tehnologije in umetna inteligenca
- Razprava

Gradiva, povezave:

<http://www.lojze.si/spela>



Globalization and
Localization Association



Reengineering the Global Content Supply Chain

Power to the Data Owners

Farewell Localization: Welcome to the
Brave New World of AI

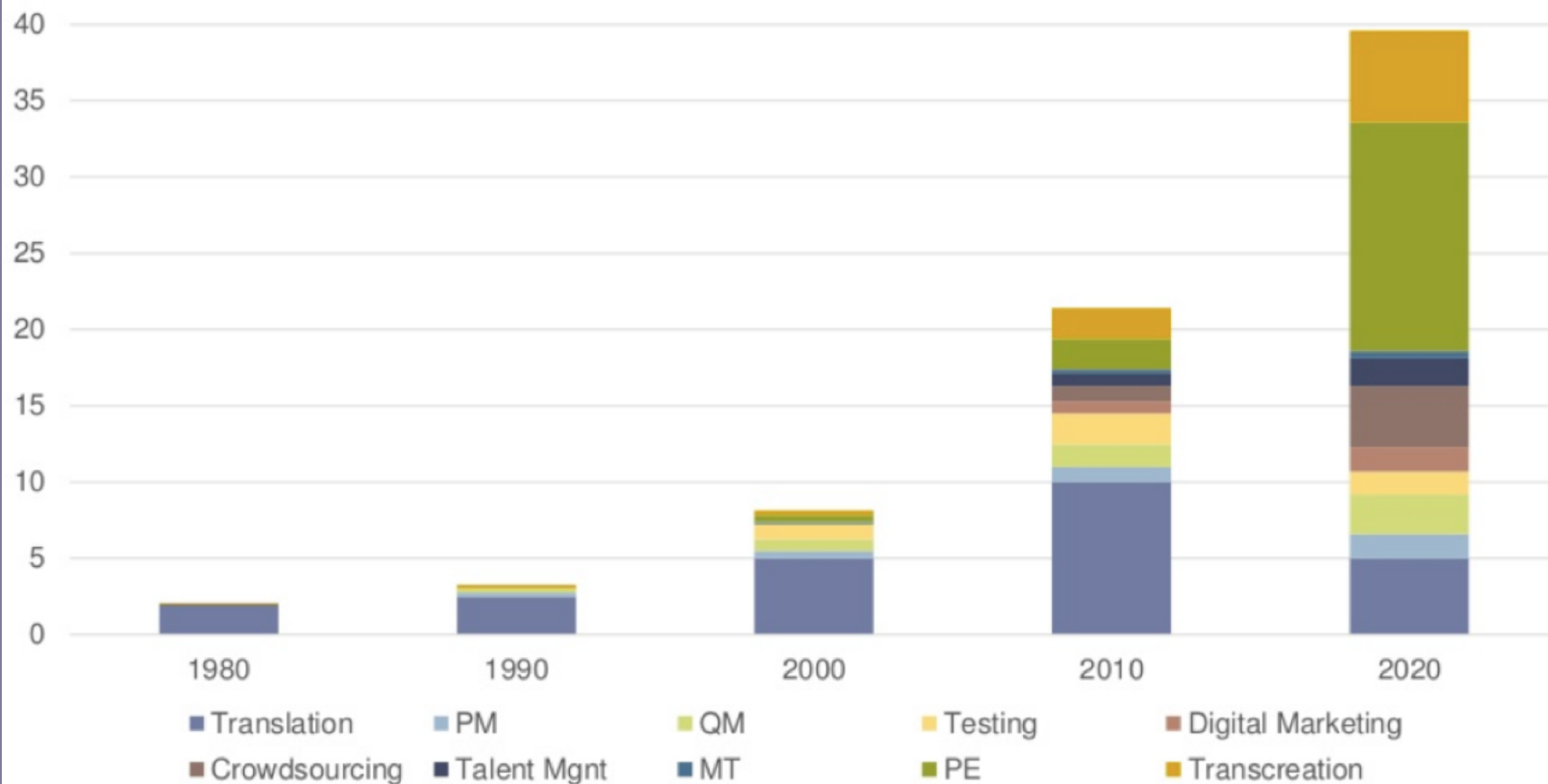
An abstract graphic with a dark brown background, featuring a network of light brown lines and nodes, resembling a neural network or a data flow diagram. The text "New Era of Artificial Intelligence" is overlaid in a bright green, sans-serif font.

New Era
of
Artificial Intelligence

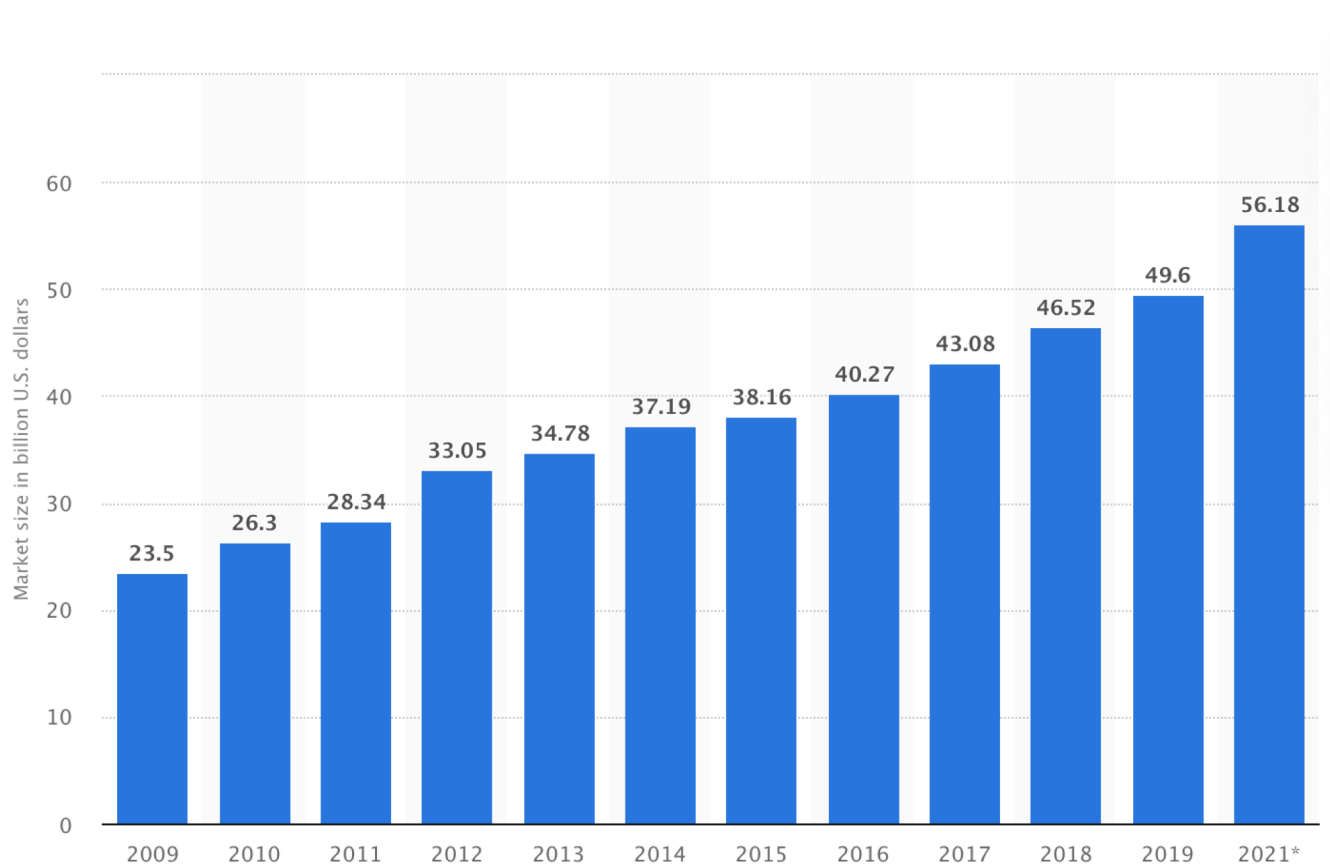
Data is the new fuel. Not just for machine learning and machine translation but also for the new extended language services: data management, data classification, cleaning, crawling, alignment, annotation and privacy.

Will Machine Learning AI Make
Human Translators An
Endangered Species?

Pestrost jezikovnih storitev

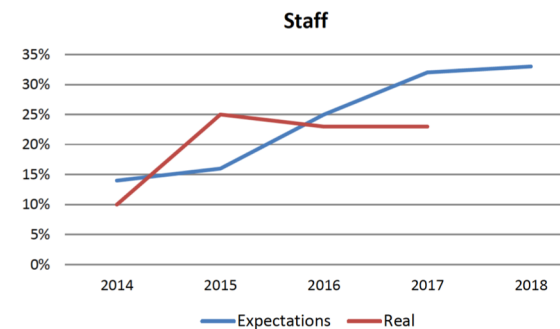
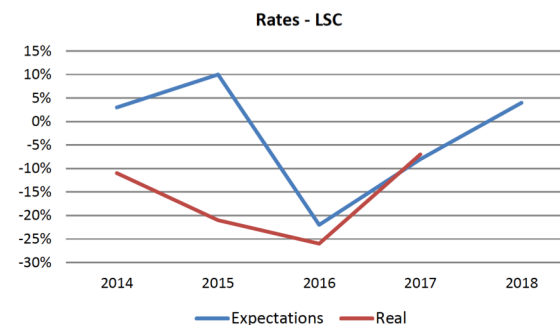
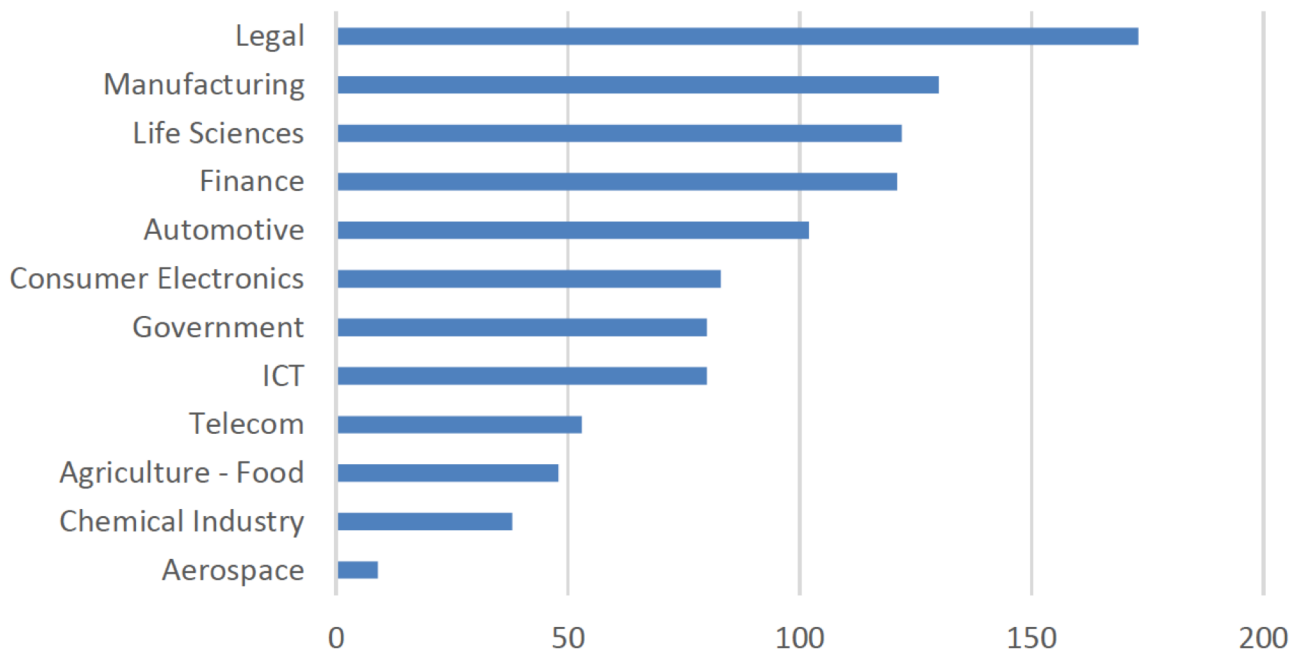


Projekcije rasti pred Covid-19



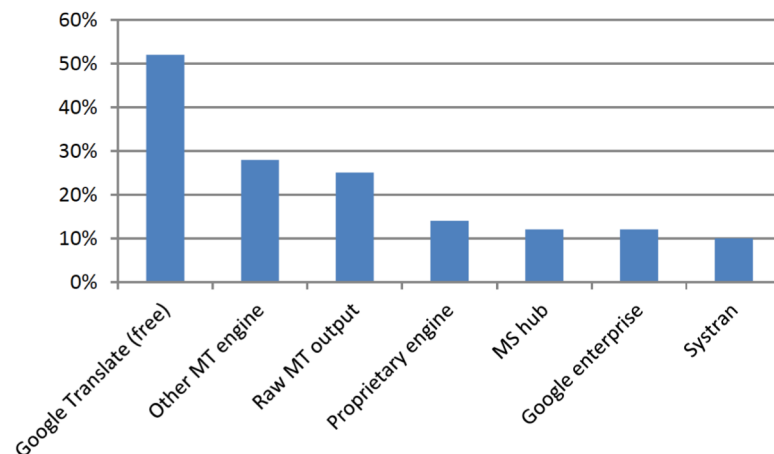
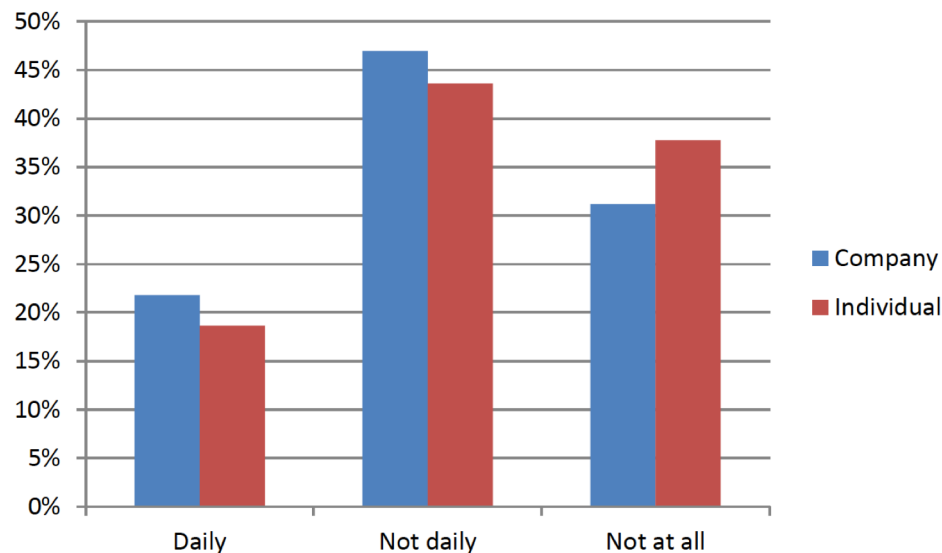
Stanje v Evropi – LIS19

Verticals companies



Stanje v Evropi – LIS19

- raba strojnih prevajalnikov



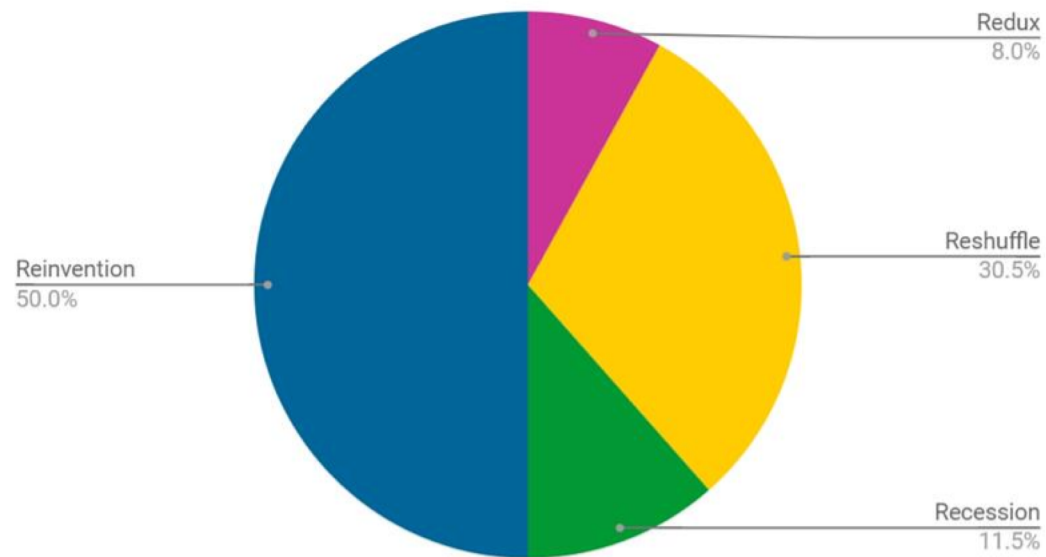
Reengineering the Global Content Supply Chain

Farewell Localization: Welcome to the

A large, bold, yellow number "4" is centered within a yellow rectangular border. The background of the slide is a dark gray with a faint, colorful illustration of interlocking puzzle pieces, a ladder, and stacks of coins.

POST-CORONA
RETHINK
SCENARIOS

State of the Translation Industry in the Summer of 2021



Prevajalstvo po koroni

- **padec naročil**, zlasti v sektorjih, povezanih s turizmom, potovanji in gostinstvom,
- **negotovost** tudi v drugih gospodarskih sektorjih,
- okrepitev trendov, ki so bili opazni že prej:
 - **tehnologije**,
 - storitve po meri -> **fleksibilnost**,
 - poslovanje na **daljavo**.
- skokovito naraščanje potreb po ponudbi **digitalnih vsebin**:
 - spletna prisotnost podjetij,
 - e-trgovina,
 - e-izobraževanje,
 - zabava;
- **naraščanje potreb** na določenih področjih:
 - medicina, farmacija, okolje, prehrana, varnost, obramba.

Prevajalstvo po koroni

- **padec naročil**, zlasti v sektorjih, povezanih s turizmom, potovanji in gostinstvom,
- **negotovost** tudi v drugih gospodarskih sektorjih,
- okrepitev trendov, ki so bili opazni že prej:
tehnologije,
 - storitve po meri -> **fleksibilnost**,
 - poslovanje na **daljavo**.
- skokovito naraščanje potreb po ponudbi **digitalnih vsebin**:
 - spletna prisotnost podjetij,
 - e-trgovina,
 - e-izobraževanje,
 - zabava;
- **naraščanje potreb** na določenih področjih:
 - medicina, farmacija, okolje, prehrana, varnost, obramba.

Prodor umetne inteligence

- nevronske strojno prevajanje (NMT)
- vse več akterjev:
 - Google, Microsoft, Yandex, Amazon, Baidu, DeepL, Tilde, ModernMT, IBM, SDL ...
 - sistemi po meri: KantanMT, Omniscien + vsi zgornji
 - ponudniki agregiranih storitev (dostop do 30+ prevajalnikov prek enega posrednika): Intento, Memsource; ocena kakovosti prevajalnika (MTQE)
 - interni sistemi (TransPerfect, Lionbridge itd.)
- novosti:
 - prevajanje na ravni dokumenta
 - sprotno prilagajanje uporabnikovemu slogu / terminologiji
 - aktivno učenje iz popravkov

Moč podatkov

- upravljanje podatkov: prioriteta ponudnikov jezikovnih storitev; podatki so kapital
- kakšni podatki?
 - dvojezični, področno-specifični
 - terminologija, slog, “prepoznavni ton podjetja”
 - enojezični, specializirani
 - gradnja jezikovnih rešitev po meri
 - pogovorni sistemi, pomoč uporabnikom, razvrščanje itd.
 - metapodatki
 - področje, jezikovni par, obseg, čas; analiza popravkov, zadetkov, napak, oznak; uporabljene tehnologije
 - za lažjo avtomatizacijo poslovnih procesov
- pojav podatkovnih tržnic
- podatkovna analitika

Strojno prevajanje in slovenščina

- Amebis Presis RBMT 2002/2003
- Google SMT 2006, NMT 2016/17
- MT@EC SMT 2013, NMT 2018
- Microsoft Bing SMT 2007, NMT 2018

[Domov](#) > [Zbirke](#) > [Javne objave](#) >

Javni razpis »Razvoj slovenščine v digitalnem okolju – jezikovni viri in tehnologije«

Predmet javnega razpisa je sofinanciranje izvajanja večletne operacije za gradnjo jezikovnih virov in tehnologij na področjih: vzdrževanje in nadgradnja korpusov, govorne tehnologije, semantični viri in tehnologije, strojno prevajanje, terminološki portal ter vzdrževanje infrastrukturnega centra za jezikovne vire in tehnologije.

Rok za oddajo vlog na javni razpis smo podaljšali do 12. februarja 2020.

ŠIFRA

JR-ESRR-Razvoj slovenščine v digitalnem okolju

DATUM OBJAVE

29. 11. 2019

ROK ZA PRIJAVO

12. 2. 2020 do 23.59

INSTITUCIJA

Ministrstvo za kulturo

Prevajalska orodja

Prevajalska orodja - lestvica

- SDL Studio (67 %)
- memoQ (49 %)
- Memsource (~ <25 %)
- Across (~ <20 %)
- Wordfast (~ <20 %)

Drugi oblačni: Transifex (L10N), SmartCat, MateCat, Wordfast Anywhere

Prevajalska orodja

- Prevajalska namizja so postala prekompleksna za uporabnike.
 - **poenostavljanje** uporabniških vmesnikov
 - perspektiva: **prilagajanje** uporabnikom z metodami UI (po zgledu Adobe)
- Prevajanje se seli v **oblak**.
 - programje kot storitev (SaaS); različne naročniške sheme
 - učinkovita avtomatizacija in optimizacija rutinskih opravil
 - vse manj pomislekov glede varnosti (pa vendar popoln prehod v oblak še ni na vidiku)
- **Integracija** prevajalnikov in drugih eksternih orodij
- Integracija **govornih** tehnologij.

Demo

- MateCat – odprtokodno brezplačno prevajalsko okolje v oblaku
- Memsources – plačljivo večuporabniško okolje za prevajanje in vodenje projektov v oblaku (trenutno vodilni na trgu)
- SmartCat – plačljivo večuporabniško okolje za prevajanje in vodenje projektov v oblaku