



LEĻMŌ IĻĻISKŪOL
LĪVŌD
INSTITŪT



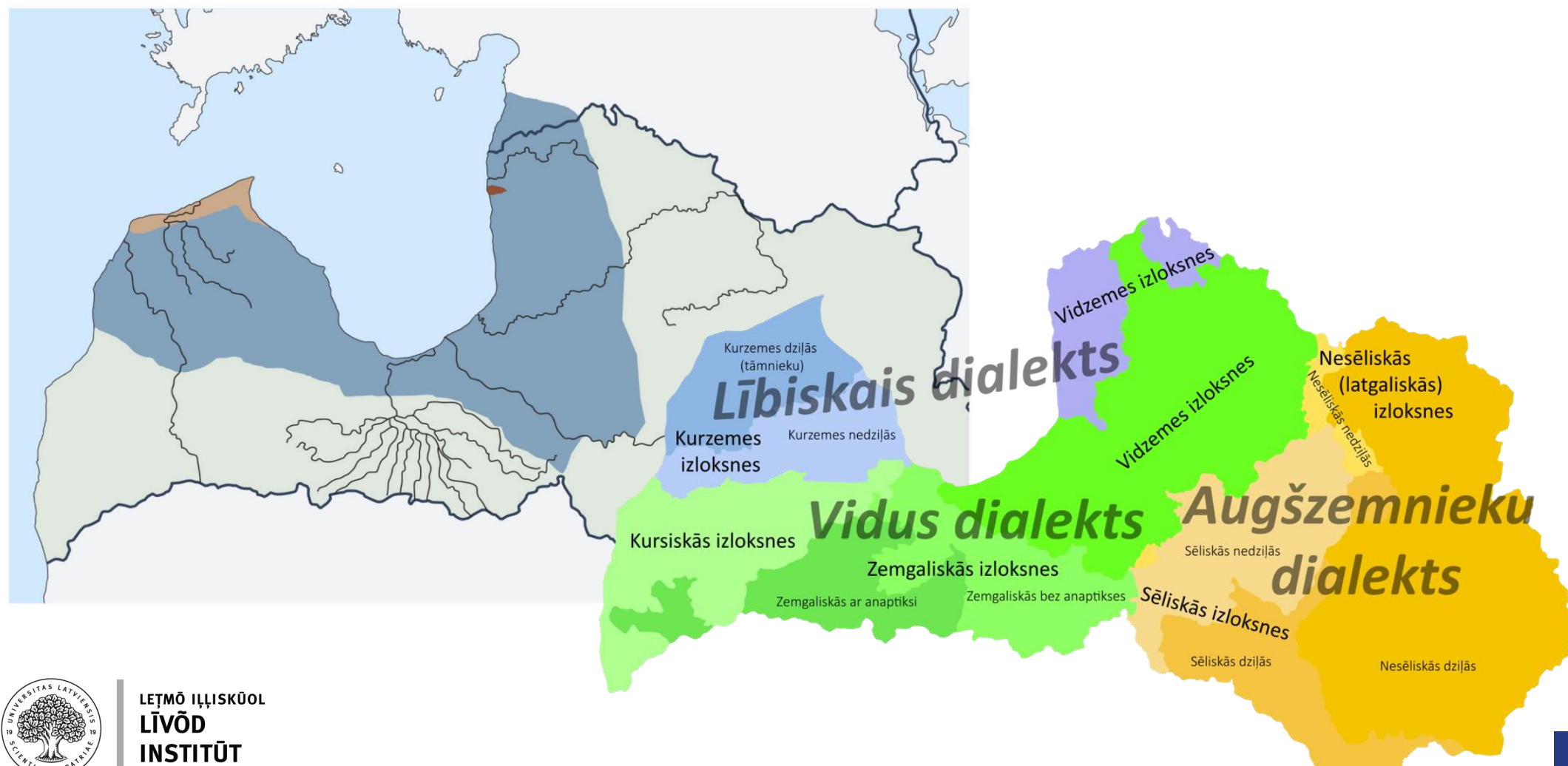
Building the Future Diversity: the University of Latvia Livonian Institute

Valts Ernštreits, UL Livonian Institute

National Research Programme “Digital Resources of the Humanities” Towards Development Of Open
And Fair Digital Humanities Ecosystem In Latvia (Dheli, 2022-2025; VPP-IZM-DH-2022/1-0002)



Livonians – who are they?



Is it necessary to pay? I doubt it.



Ar reikia mokėti? Aš labai abejoju.



Kas on vaja maksta? Ma väga kahtlen.



Vai vajag maksāt? Es ļoti šaubos.



Voi väga maksuot? As cīši šaubūs.



Või um vajāg maksõ? Ma vāggi kōdlōb.



Livonians – who are they?

- **Indigenous people of Latvia since 1991 with a long history and great importance in the Baltic region**
- Currently, there are **around 1000–2000 Livonians, ca 20 fluent speakers**, the growing rate of interest among the younger generation (over 50 children aged 7–17 are practicing Livonian culture annually)
- **Community is scattered** across Latvia due to both World Wars and subsequent Soviet occupation, **it is weak, fragmented, and extremely diverse**, but at the same time – **unproportionally active** in sustaining its identity, safeguarding and developing its modern culture



The UL Livonian Institute

- **Established in September 2018, built from scratch as a startup based on projects; research in three directions:**
 - 1) Research of the Livonians** – past 200 years of language, tangible and intangible heritage, culture, social structures, etc;
 - 2) Research of the Livonian heritage** – Livonian traces in Latvian language, culture, and history; working closely with the general public
 - 3) Livonian and global experience** – strategies and methods for the sustainability of endangered communities



LEŅMŌ IĻĪSKŪOL
LĪVŌD
INSTITŪT



VPP
Valsts pētījumu
programma



Key tool – Livonian.tech

- **Vocabulary** – words and phrases with translations
- **Grammatical data** – forms, variations
- **Textual data** – corpora, texts in the Livonian with translations
- **Audiovisual data** – words and texts voiced by Livonians
- **Sources / people** – information on sources and people behind them
- **Geospatial data** – locations, place names etc
- **Cultural data** – information on events, stories, physical and digital objects



Sound

Scroll to top



būom

beam, boom



doņņõ

to be naughty, to throw a
tantrum, to be capricious



T̄d | vastālit

equality



agā

perhaps, maybe

Follow Us
f -

Livonian ▾



Enter keyword...

Search



Vārdi Formas Teksti

Karte



Vietas



Valoda LV EE EN Kārtoš

Redzami 1 - 1 no 1 šķirklīem



Redzami 1 - 1 no 1 šķirklīem



Meklēt



Skaits 50

Pievienot

< 1 >



Pievienot

< 1 >



Petš
Niklon



Pizā,
1912

Pizā / *25
Pizā □ Miķeļtornis, Pize

Place
name

Sīlda kārand



Legend

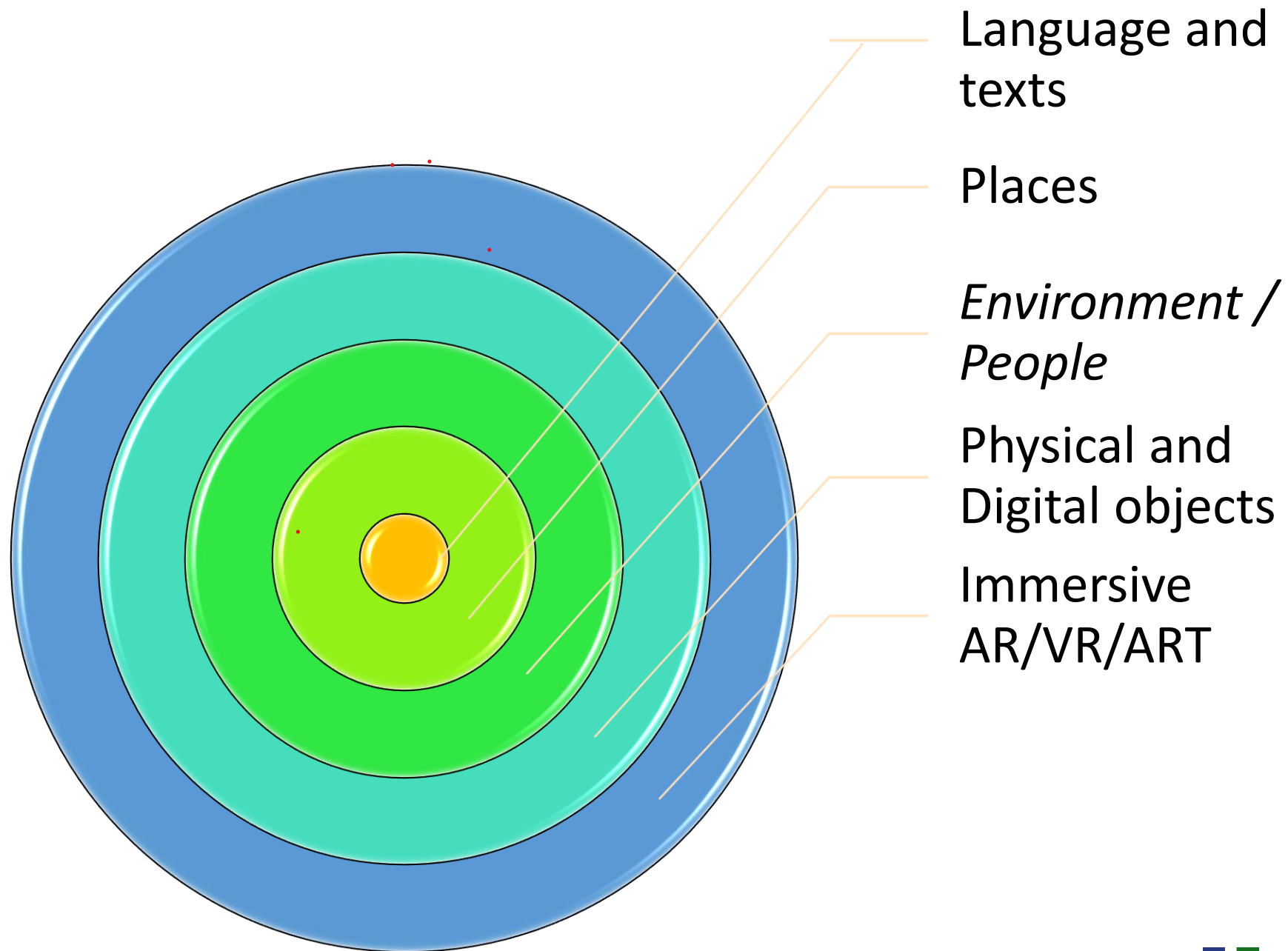
*Pr 4. Nēstst kū pāl.
Frš nēstst um lānd, kas vjed-pānš
kāids kouvn-aigš. Un nē ta um nā'gtēn
sormērs kū pāl: "Van kū!" Un nē se
kū um vētēn tām entē šī 77 5l amad
pānš dēvs.*

Item



pīp s *122
piip □ pīpe
◆ **pīpō tū'gtō** piipu pahvida □ kūpināt pīpi.

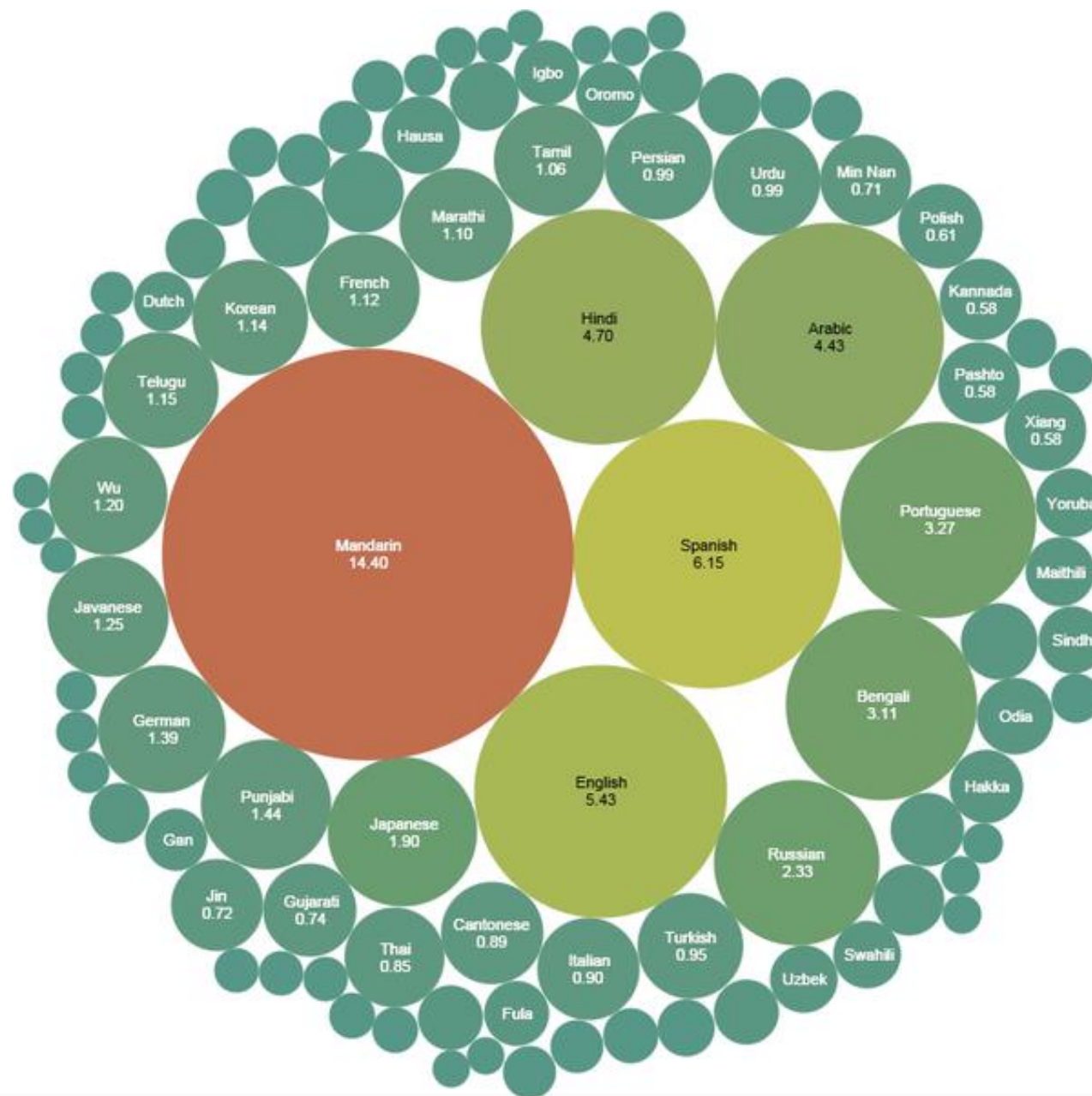




Why Livonians ?

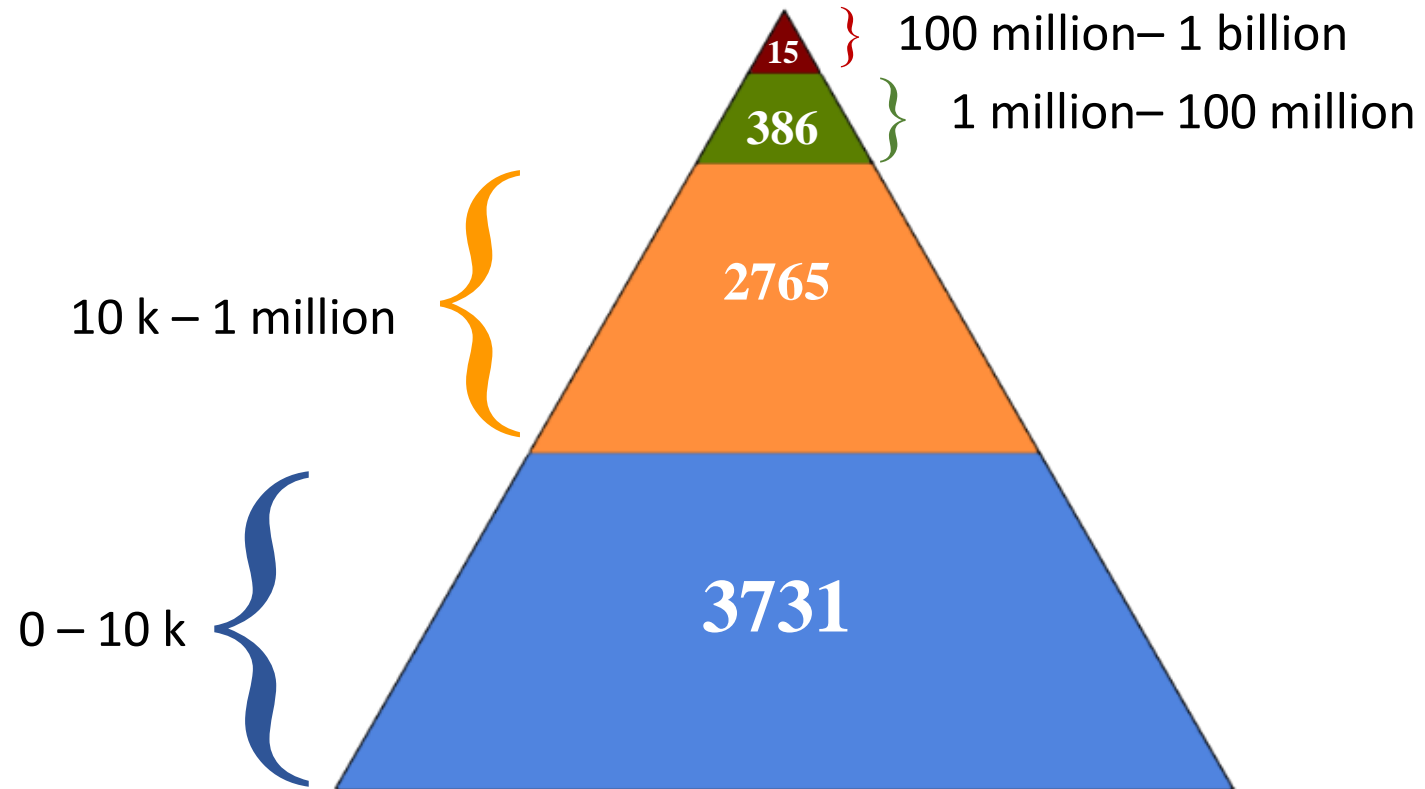
- Working with Livonians we are exploring the future as Livonians have been virtual for the past 80+ years
- Livonians have developed some kind of resistance sustaining the community, identity, and ICH
- Processes are compactly, but relatively well-documented





LETMŌ IĻĪSKŪOL
LĪVŌD
INSTITŪT





LEŅMŌ IĻĪSKŪOL
LĪVŌD
INSTITŪT



Our Goals

- **Sustainability** of critically endangered languages and cultures
- Knowledge about **processes in fragmented communities at the doorstep of virtualization** in the global information and AI society
- Cultural and linguistic **diversity of the humankind**



LEŅMŌ IĻĪSKŪOL
LĪVŌD
INSTITŪT



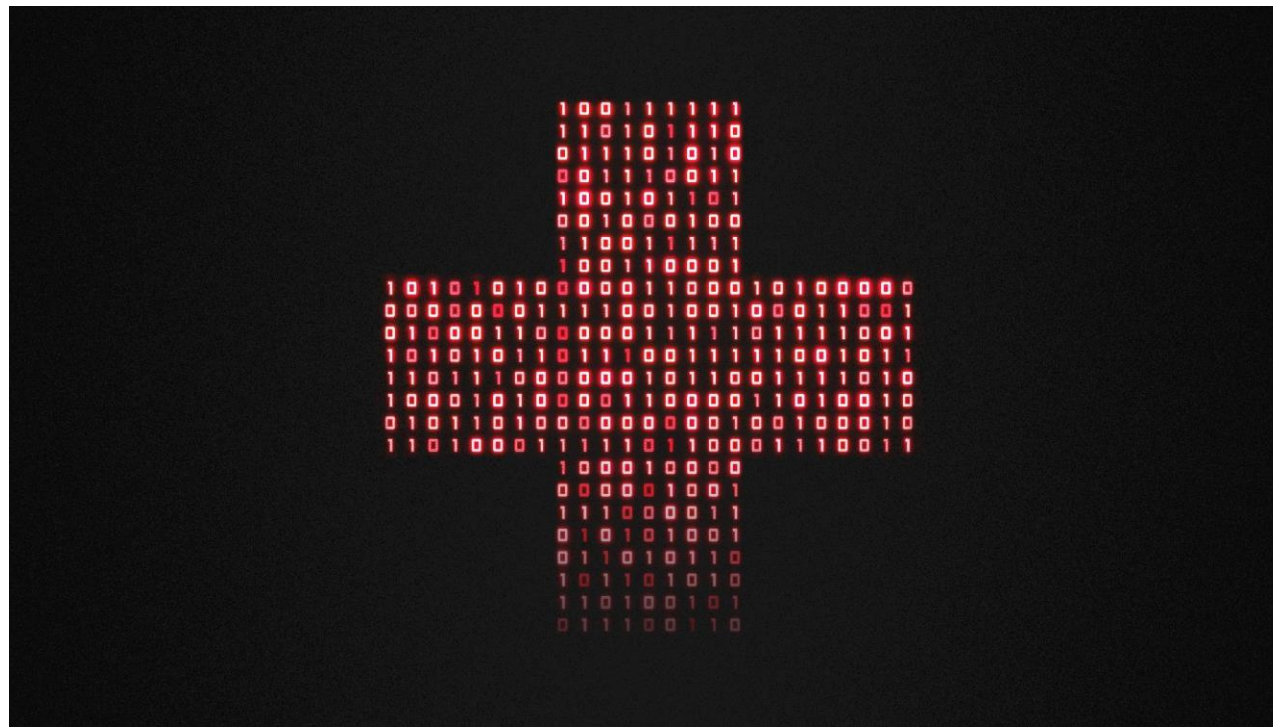
Challenges

- **Under-researched** – lack of knowledge, incorrect interpretation of processes, lack of right perspective of tools needed and feasible
- **Under-documented** – fragmented and limited data for all purposes including AI tools; **limited data, but the same results to be achieved**
- **Under-resourced** – lack of financial and human resources to perform necessary actions
- **Diversity** – every situation is unique, there are no transferrable sets, just strategies, approaches, methods, and tools



Benefits

- **Limited data – better overview** and somewhat easier data handling, less computing capacities
- **The key benefit – data integration and cross-usage**



Things to remember

- Protection of the community – **no data processing or products should be harmful to the community**
- **Tech to support people** – semi-automated approaches, augmented Golden Standard is the key to the quality when resources are low
- **Zero Waste policy** – all efforts should be meaningful, no data or resource is useless
- **Dreams, not fantasies** – resources to be planned based on need and effort necessary



Strategies: big data *versus* small data



LEŅMĀ IZĻĪSKŪOL
LĪVŌD
INSTITŪT



Strategies: statistics and approaches

- **Limited opportunities for statistical decisions** – for big data statistics erase deviations (1 out of 1000000 is a deviation), small data – 1 out of 10 may be correct; **quality methods developed**
- **Flexible approaches, inclusive standards, and skill-based task division** – tools are planned to provide involvement of people with different levels of skills; **reactive products and developments** to obtain new data (e.g. cultural data from geospatial activities)
- **Open and unfinished:** platform-free online access, only necessary interpretations and decisions according to existing knowledge, several options – **gaps left open to avoid being stuck**



Strategies: planning and applications

- **Rule-based:** simple algorithms to support speeding up manual annotation, the inclusion of statistical data and hypothetical data, building of declension paradigms, spellchecking, helping to ensure documentation consistency (through other languages); mapping
- **AI-based:** machine translation (using other languages) to provide accessibility to the ICH; speech synthesis and speech recognition (double-trained AI) providing access to the content
- **Mixed applications:** machine translation, form prediction, language acquisition, subtitling etc



Our friends



UPPSALA
UNIVERSITET



INDIANA UNIVERSITY
BLOOMINGTON

TE TAURA WHIRI
I TE REO MĀORI
MĀORI LANGUAGE COMMISSION



LAMI LU LITERATŪRAS,
FOLKLORAS
UN MĀKSLAS
INSTITŪTS



university of
 groningen



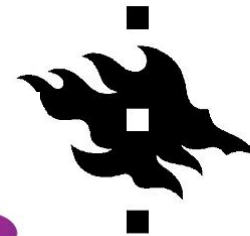
UNIVERSITY OF
EASTERN FINLAND



KULTUURIMINISTEERIUM



Eesti Geenivaramu



UNIVERSITY OF HELSINKI

SUOMALAIS-UGRILAINEN SEURA



Eesti Kirjandusmuuseum



LETMÕ IĻĪSKŪOL
LĪVŌD
INSTITŪT





2022-2032 | INTERNATIONAL DECADE OF

Indigenous Languages



LĪVŌD
PIERĀNDŌKS ĀIGAST
2023

A stylized blue fish logo with three green circles on its head, positioned above the text.



LEĶMŌ IĻĻISKŪOL
LĪVŌD
INSTITŪT



Thank You! | Tienū!

