EUROPEAN LANGUAGE EQUALITY

META-NET
META-FORUM 2022

EUROPEAN LANGUAGE GRID

Language Technology Vision for 2030

Jan Hajič (Charles University), Georg Rehm (DFKI)
hajic@ufal.mff.cuni.cz, georg.rehm@dfki.de

08/09-06-2022 META-FORUM 2022 – Joining the European Language Grid (hybrid conference)
http://www.european-language-grid.eu – http://european-language-equality.eu

Session 10: Language Technology Vision for 2030

14:00 Our Vision for 2030: Digital Language Equality – Jan Hajič (Charles University, Czech Republic)

14:10 Strategic Research and Innovation Agenda – Georg Rehm (DFKI, Germany)

14:30 Discussion

Our Vision for 2030: Digital Language Equality

Digital Language Equality by Language Technology – the big picture

Deep Dives
(prev. section)

SoA, Gaps

Vision

Recommendations

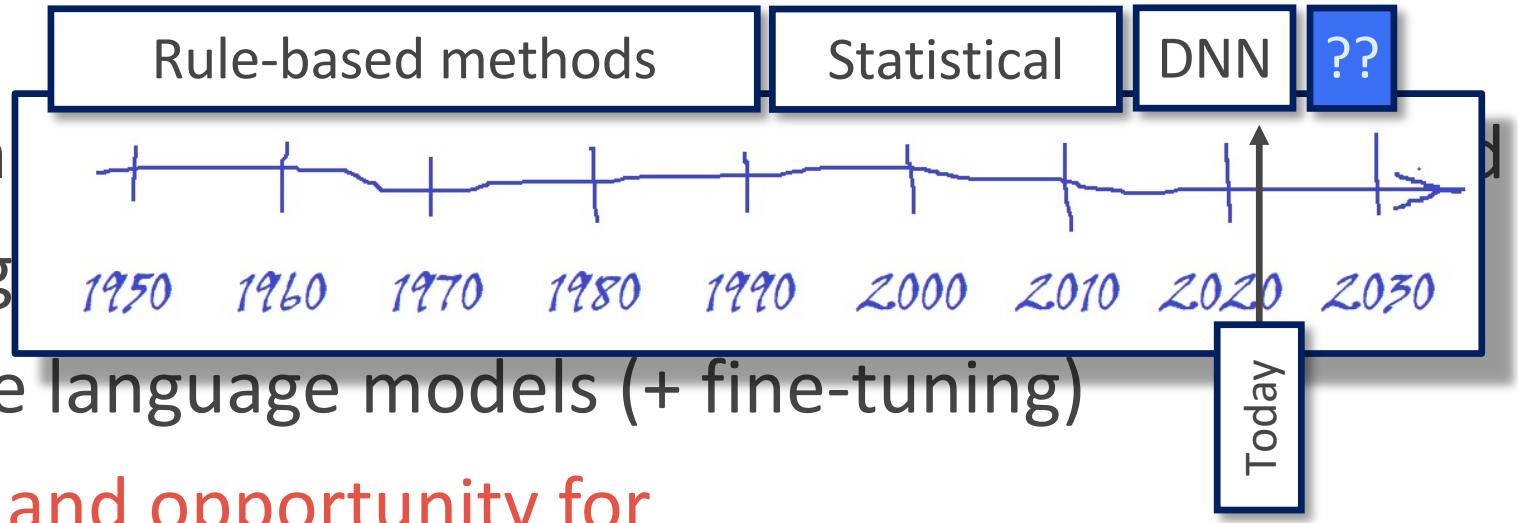
Surveys

Language reports

Interviews

Our Vision for 2030: Digital Language Equality

- The facts
  - Paradigm
  - Multilang
  - Very large language models (+ fine-tuning)



learning

- Challenge and opportunity for
  - Low-resourced languages (→ Digital Language Equality)
  - New applications (by increasing quality or the underlying LT)
  - Natural Language Understanding (and “General AI”)

Our Vision for 2030: Language Technology Landscape

- Goal: Natural Language Understanding in/for General AI
  - Priority areas
    - Machine Translation
      - Transparent, context- and culture-sensitive ubiquitous MT, incl. spoken translation
    - Speech Technologies
      - Robust Speech UI integrated to any application in any environment; incl. sign languages
    - Text processing and analytics
      - Human-in-the-loop, (un)structured data integration, 200 (500?) languages
    - Data and Knowledge bases
      - Data in all EU official, regional and other languages in repositories
      - “FAIR” and part of EOSC, abundant, for E2E systems, commercially usable
      - Knowledge bases linked to all languages

Our Vision for 2030: Impact on Society

- Transparent, barrier-free communication and mobility across Europe
  - Among citizens, businesses and governments/administrations
- Education
  - Language teaching, cultural exchange, access to foreign culture and cultural heritage
- Language content creation
  - Including multimodal production (images, A/V, …)
- Healthcare applications, from diagnostics to elderly care
  - Cross-border healthcare availability

Our Vision for 2030: Not throwing everything out …

- Strong research, including basic research on ML, methodology, evaluation, …
- Progress through competition in applied research/engineering
  - Shared tasks, datasets, evaluation platforms
- Data collection
  - New data for novel tasks, data procurement, data cleaning, more data for identified languages
  - Integration of symbolic databases and speech/textual/multimodal data
- Higher Education in LT/AI/ML, including PhD-level, talent retention & acquisition
  - Student participation in EU project (such as HE, 10th FP, …)
- Fight on the legal front
  - Data openness and usability, ethical issues, bias avoidance, level field across the world
- Public support
  - For most of the above areas

Session 10: Language Technology Vision for 2030

14:00 Our Vision for 2030: Digital Language Equality – Jan Hajič (Charles University, Czech Republic)

14:10 Strategic Research and Innovation Agenda – Georg Rehm (DFKI, Germany)

14:30 Discussion

Strategic Research and Innovation Agenda

EUROPEAN LANGUAGE EQUALITY

EUROPEAN LANGUAGE GRID

Consortium: 52 partners from all over Europe

Coordinator: ADAPT Centre (Dublin City University)

Co-Coordinator: DFKI

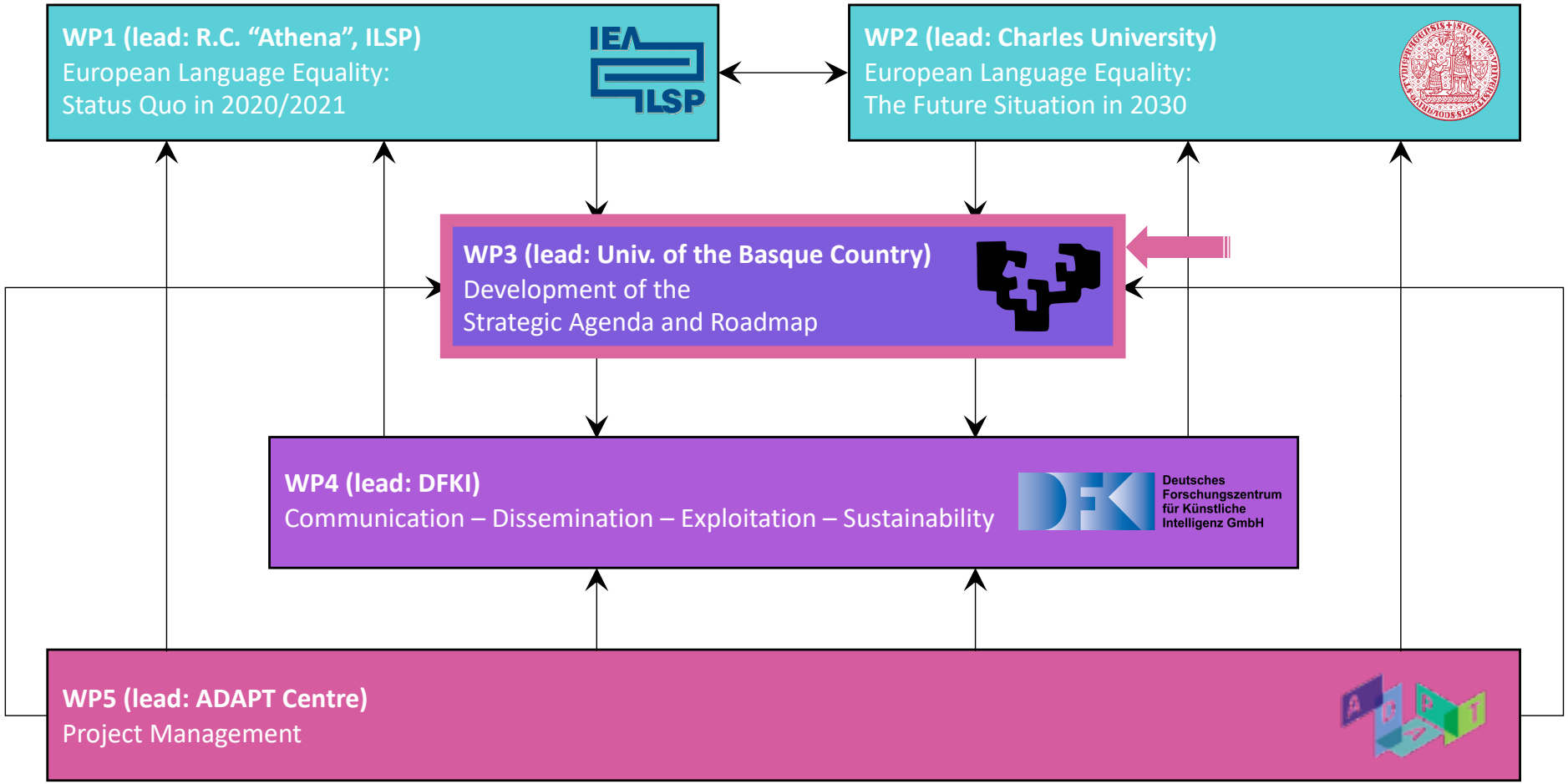
Objective: ***development of a strategic research, innovation and implementation agenda to achieve digital language equality in Europe by 2030***

Runtime: 18 months – ELE and ELG are both finishing up in June 2022

Started on **1 January 2021**

http://www.european-language-equality.eu

EUROPEAN LANGUAGE EQUALITY



Work Packages

EUROPEAN LANGUAGE EQUALITY

WP1 European Language Equality: Status Quo in 2020/2021
Task 1.1: Defining Digital Language Equality
Task 1.2: Language Technologies and Language-centric AI – State of the Art
Task 1.3: Language Technology Support of Europe's Languages in 2020/2021

WP2 European Language Equality: The Future Situation in 2030
Task 2.1: The perspective of European LT developers (industry and research)
Task 2.2: The perspective of European LT users and consumers
Task 2.3: Science – Technology – Society: Language Technology in 2030

- Digital Language Equality: Definition of the concept
- Language Technology and language-centric AI: State of the Art
- 32 reports on the technology support of 32 European languages (META-NET White Paper update)
- Reports from networks, initiatives and associations
- Deep dives (MT, speech, text analytics, data)
- Report on external consultations and surveys
- Forecast: Language Technology in 2030

WP3 Development of the Strategic Agenda and Roadmap
Task 3.1: Desk research – landscaping
Task 3.2: Consolidation and aggregation of all input received
Task 3.3: Final round of feedback collection

- Existing strategic documents and projects in LT/AI
- Strategic agenda and roadmap: initial version
- Final round of feedback collection
- Strategic agenda and roadmap: final version

WP4 Communication – Dissemination – Exploitation – Sustainability
Task 4.1: Overall project communication and dissemination
Task 4.2: Liaise with EP/EC – organisation of a targeted workshop
Task 4.3: Organisation of final ELE conference
Task 4.4: Production of PR materials and sustainable results

- EP/EC Workshop → ELE Conference
- ELE Strategic Agenda and Roadmap (print version, interactive version) → Final ELE Book Publication

WP5 Project Management
Task 5.1: Overall project management including Project Management Office
Task 5.2: Digital collaboration and document management infrastructure

SRIA

No.	Deliverable name	WP	Short name	Type	Diss. level	Date
D1.1	Digital Language Equality – preliminary definition	1	DCU	R	Public	3
D1.2	Report on the state of the art in Language Technology and Language-centric AI	1	EHU	R	Public	9
D1.3	Digital Language Equality – full specification of the concept	1	DCU	R	Public	13
D1.4	Report on Basque	1	EHU	R+OTH	Public	14
D1.5	Report on Bulgarian	1	IBL	R+OTH	Public	14
D1.6	Report on Catalan	1	BSC	R+OTH	Public	14
D1.7	Report on Croatian	1	FFZG	R+OTH	Public	14
D1.8	Report on Czech	1	CUNI	R+OTH	Public	14
D1.9	Report on Danish	1	UCPH	R+OTH	Public	14
D1.10	Report on Dutch	1	UU	R+OTH	Public	14
D1.11	Report on English	1	USFD	R+OTH	Public	14
D1.12	Report on Estonian	1	UTART	R+OTH	Public	14
D1.13	Report on Finnish	1	UHEL	R+OTH	Public	14
D1.14	Report on French	1	CNRS	R+OTH	Public	14
D1.15	Report on Galician	1	UVIGO	R+OTH	Public	14
D1.16	Report on German	1	DFKI	R+OTH	Public	14
D1.17	Report on Greek	1	ILSP	R+OTH	Public	14
D1.18	Report on Hungarian	1	NYTI	R+OTH	Public	14
D1.19	Report on Icelandic	1	SAM	R+OTH	Public	14
D1.20	Report on Irish	1	DCU	R+OTH	Public	14
D1.21	Report on Italian	1	FBK	R+OTH	Public	14
D1.22	Report on Latvian	1	IMCS	R+OTH	Public	14
D1.23	Report on Lithuanian	1	LKI	R+OTH	Public	14
D1.24	Report on Luxembourgish	1	LIST	R+OTH	Public	14
D1.25	Report on Maltese	1	UOM	R+OTH	Public	14
D1.26	Report on Norwegian	1	LCNOR	R+OTH	Public	14
D1.27	Report on Polish	1	PAS	R+OTH	Public	14
D1.28	Report on Portuguese	1	ULIS	R+OTH	Public	14
D1.29	Report on Romanian	1	ICIA	R+OTH	Public	14
D1.30	Report on Serbian	1	FILFUB	R+OTH	Public	14
D1.31	Report on Slovak	1	JULS	R+OTH	Public	14
D1.32	Report on Slovenian	1	JSI	R+OTH	Public	14
D1.33	Report on Spanish	1	BSC	R+OTH	Public	14
D1.34	Report on Swedish	1	KTH	R+OTH	Public	14
D1.35	Report on Welsh	1	BNGR	R+OTH	Public	14
D1.36	Database and dashboard with the empirical data collected in D1.4-D1.35 (and others)	1	ILSP	R+OTH	Public	16
D2.1	Specification of the consultation process including templates, surveys, events etc.	2	CUNI	R	Public	4
D2.2	Report from CLAIRE	2	ULEID	R+OTH	Public	14
D2.3	Report from CLARIN	2	CLARIN	R+OTH	Public	14
D2.4	Report from LT Innovate	2	CRSLNG	R+OTH	Public	14
D2.5	Report from META-NET	2	CUNI	R+OTH	Public	14
D2.6	Report from ELG	2	DFKI	R+OTH	Public	14
D2.7	Report from ECSPM	2	ECSPM	R+OTH	Public	14
D2.8	Report from EFNIL	2	EFNIL	R+OTH	Public	14
D2.9	Report from ELEN	2	ELEN	R+OTH	Public	14
D2.10	Report from LIBER	2	LIBER	R+OTH	Public	14
D2.11	Report from NEM	2	ERSCM	R+OTH	Public	14
D2.12	Report from Wikipedia	2	WMD	R+OTH	Public	14
D2.13	Technology deep dive Machine Translation	2	TILDE	R	Public	14
D2.14	Technology deep dive Speech Technologies	2	SAIL	R	Public	14
D2.15	Technology deep dive Text Analytics and Natural Language Understanding	2	EXPSYS	R	Public	14
D2.16	Technology deep dive Data	2	SWC	R	Public	14
D2.17	Report on all external consultations and surveys	2	CUNI	R	Public	15
D2.18	Report on the state of Language Technology in 2030	2	CUNI	R	Public	16
D3.1	Report on existing strategic documents and projects in LT/AI	3	EHU	R	Public	3
D3.2	Strategic agenda including roadmap – initial version	3	DFKI	R	Public	15
D3.3	Report on the final round of feedback collection	3	EHU	R	Public	17
D3.4	Strategic agenda including roadmap – final version	3	DFKI	R	Public	18
D4.1	Promotional materials and PR package	4	DFKI	R+OTH	Public	3
D4.2	Communication and dissemination plan	4	DFKI	R	Public	6
D4.3	Report on EP/EC workshop	4	DFKI	R	Public	18
D4.4	Report on ELE conference	4	DFKI	R	Public	18
D4.5	Strategic agenda and roadmap (print version, online version)	4	DFKI	R	Public	18
D4.6	ELE book publication	4	DFKI	R	Public	18
D5.1	Digital collaboration and document management infrastructure	5	DCU	R+OTH	Confidential	3
D5.2	Project management report (interim report)	5	DCU	R	Confidential	9
D5.3	Project management report (final report)	5	DCU	R	Confidential	18

EUROPEAN LANGUAGE EQUALITY

Deliverables

No.	Deliverable name	WP	Short name	Type	Diss. level	Date
D1.1	Digital Language Equality – preliminary definition	1	DCU	R	Public	3
D1.2	Report on the state of the art in Language Technology and Language-centric AI	1	EHU	R	Public	9
D1.3	Digital Language Equality – full specification of the concept	1	DCU	R	Public	13
D1.4–D1.35	Report on Basque ... Report Welsh	1	EHU	R+OTH	Public	14
D1.36	Database and dashboard with the empirical data collected in D1.4-D1.35 (and others)	1	ILSP	R+OTH	Public	16
D2.1	Specification of the consultation process including templates, surveys, events etc.	2	CUNI	R	Public	4
D2.2–D2.12	Report from CLAIRE ... Report from Wikipedia	2	ULEID	R+OTH	Public	14
D2.13–D2.16	Technology deep dives MT, Speech, Text Analytics, Data	2	TILDE	R	Public	14
D2.17	Report on all external consultations and surveys	2	CUNI	R	Public	15
D2.18	Report on the state of Language Technology in 2030	2	CUNI	R	Public	16
D3.1	Report on existing strategic documents and projects in LT/AI	3	EHU	R	Public	3
D3.2	Strategic agenda including roadmap – initial version	3	DFKI	R	Public	15
D3.3	Report on the final round of feedback collection	3	EHU	R	Public	17
D3.4	Strategic agenda including roadmap – final version	3	DFKI	R	Public	18
D4.1	Promotional materials and PR package	4	DFKI	R+OTH	Public	3
D4.2	Communication and dissemination plan	4	DFKI	R	Public	6
D4.3	Report on EP/EC workshop	4	DFKI	R	Public	18
D4.4	Report on ELE conference	4	DFKI	R	Public	18
D4.5	Strategic agenda and roadmap (print version, online version)	4	DFKI	R	Public	18
D4.6	ELE book publication	4	DFKI	R	Public	18
D5.1	Digital collaboration and document management infrastructure	5	DCU	R+OTH	Confidential	3
D5.2	Project management report (interim report)	5	DCU	R	Confidential	9
D5.3	Project management report (final report)	5	DCU	R	Confidential	18

Deliverables

(condensed)

No.	Deliverable name	Short name	Date
D1.1	Digital Language Equality – preliminary definition	DCU	3
D3.1	Report on existing strategic documents and projects in LT/AI	EHU	3
D2.1	Specification of the consultation process including templates, surveys, events etc.	CUNI	4
D1.2	Report on the state of the art in Language Technology and Language-centric AI	EHU	9
D1.3	Digital Language Equality – full specification of the concept	DCU	13
D1.4–D1.35	Report on Basque ... Report Welsh	EHU	14
D2.2–D2.12	Report from CLAIRE ... Report from Wikipedia	ULEID	14
D2.13–D2.16	Technology deep dives MT, Speech, Text Analytics, Data	TILDE	14
D2.17	Report on all external consultations and surveys	CUNI	15
D3.2	Strategic agenda including roadmap – initial version	DFKI	15
D1.36	Database and dashboard with the empirical data collected in D1.4-D1.35 (and others)	ILSP	16
D2.18	Report on the state of Language Technology in 2030	CUNI	16
D3.3	Report on the final round of feedback collection	EHU	17
D3.4	Strategic agenda including roadmap – final version	DFKI	18
D4.5	Strategic agenda and roadmap (print version, online version)	DFKI	18
D4.6	ELE book publication	DFKI	18

Deliverables (condensed even further to those that are relevant for the SRIIA development)

SRIA: Outline

Strategic Research, Innovation and Implementation Agenda and a Roadmap for Achieving full Digital Language Equality in Europe by 2030

1. Multilingual Europe and Digital Technologies
2. Trends and Mega-Trends in Digital Technologies
3. Language Technology and Language-Centric Artificial Intelligence
4. Language Technology and Digital Language Equality in 2022
5. Language Technology and Digital Language Equality in 2030: The ELE Technology Vision and Priority Research Themes
6. A Shared European Programme for Language Technology and Digital Language Equality in Europe by 2030: Recommendations
7. Roadmap towards Digital Language Equality in Europe by 2030

SRIA: Outline

Strategic Research, Innovation and Implementation Agenda and a Roadmap for Achieving full Digital Language Equality in Europe by 2030

1.	Multilingual Europe and Digital Technologies	*10-16 pages*
2.	Trends and Mega-Trends in Digital Technologies	*06-10 pages*
3.	Language Technology and Language-Centric Artificial Intelligence	*06-16 pages*
4.	Language Technology and Digital Language Equality in 2022	*16-22 pages*
5.	Language Technology and Digital Language Equality in 2030: The ELE Technology Vision and Priority Research Themes	*24-30 pages*
6.	A Shared European Programme for Language Technology and Digital Language Equality in Europe by 2030: Recommendations	*12-16 pages*
7.	Roadmap towards Digital Language Equality in Europe by 2030	*04-06 pages*
		78-116 pages

SRIA: Outline

Strategic Research, Innovation and Implementation Agenda and a Roadmap for Achieving full Digital Language Equality in Europe by 2030

1.	Multilingual Europe and Digital Technologies	*10-16 pages*
2.	Trends and Mega-Tre…ds in Digital Technologies	*…06-10 pages*
3.	Language Technology…	*…06-16 pages*
4.	Language Technology…	*…16-22 pages*
5.	Language Technology… The ELE Technology Vision and Priority Research Themes	*…24-30 pages*
6.	A Shared European Programme for Language Technology and Digital Language Equality in Europe by 2030: Recommendations	*12-16 pages*
7.	Roadmap towards Digital Language Equality in Europe by 2030	*04-06 pages*
		78-116 pages

The SRIA will be published online as a stand-alone document. In addition, an abridged version, which focuses upon the main recommendations, will be published in the ELE book.

SRIA: Outline

Strategic Research, Innovation and Implementation Agenda and a Roadmap for Achieving full Digital Language Equality in Europe by 2030

1. Multilingual Europe and Digital Technologies
2. Trends and Mega-Trends in Digital Technologies
3. Language Technology and Language-Centric Artificial Intelligence
4. Language Technology and Digital Language Equality in 2022
5. Language Technology and Digital Language Equality in 2030: The ELE Technology Vision and Priority Research Themes
6. A Shared European Programme for Language Technology and Digital Language Equality in Europe by 2030: Recommendations
7. Roadmap towards Digital Language Equality in Europe by 2030

Will be mainly sourced and informed by re-using and summarising existing ELE deliverables

SRIA: Outline

Strategic Research, Innovation and Implementation Agenda and a Roadmap for Achieving full Digital Language Equality in Europe by 2030

1. Multilingual Europe and Digital Technologies
2. Trends and Mega-Trends in Digital Technologies
3. Language Technology and Language-Centric Artificial Intelligence
4. Language Technology and Digital Language Equality in 2022
5. Language Technology and Digital Language Equality in 2030: The ELE Technology Vision and Priority Research Themes
6. A Shared European Programme for Language Technology and Digital Language Equality in Europe by 2030: Recommendations
7. Roadmap towards Digital Language Equality in Europe by 2030

Will illustrate the timeline and temporal order of our recommendations

SRIA: Outline

Strategic Research, Innovation and Implementation Agenda and a Roadmap for Achieving full Digital Language Equality in Europe by 2030

1. Multilingual Europe and Digital Technologies
2. Trends and Mega-Trends in Digital Technologies
3. Language Technology and Language-Centric Artificial Intelligence
4. Language Technology and Digital Language Equality in 2022
5. Language Technology and Digital Language Equality in 2030: The ELE Technology Vision and Priority Research Themes
6. **A Shared European Programme for Language Technology and Digital Language Equality in Europe by 2030: Recommendations**
7. Roadmap towards Digital Language Equality in Europe by 2030

Main recommendations

SRIA: Outline

Strategic Research, Innovation and Implementation Agenda and a Roadmap for Achieving full Digital Language Equality in Europe by 2030

6. A Shared European Programme for Language Technology and Digital Language Equality in Europe by 2030: Recommendations

1. Overview and Main Concept
2. Research Recommendations
3. Technology and Data Recommendations
4. Infrastructure Recommendations
5. Policy Recommendations
6. Governance Model

Main language-independent recommendations extracted from the various ELE deliverables and feedback rounds.

With many thanks to Maria Giagkou, Annika Grützner-Zahn, Aritz Farwell and others who recently went through the 2000+ pages of ELE reports to extract recommendations!

SRIA: Outline

Strategic Research, Innovation and Implementation Agenda and a Roadmap for Achieving full Digital Language Equality in Europe by 2030

6. A Shared European Programme for Language Technology and Digital Language Equality in Europe by 2030: Recommendations

1. Overview and Main Concept
2. Research Recommendations
3. Technology and Data Recommendations
4. Infrastructure Recommendations
5. Policy Recommendations
6. Governance Model

The whole community contributed to the ELE deliverables, *especially through the various ELE surveys amongst the members of the included networks (CLAIRE, CLARIN, LT Innovate, META-NET, ELG, ECSPM, EFNIL, ELEN, LIBER, NEM, Wikipedia) but also through 60+ interviews with experts from the field. We estimate that 500-750 colleagues have contributed to the process.*

SRIA: Section 6.2 – Research Recommendations 1/3 *(draft)*

- **Refocus and massively strengthen European LT/NLP research through a large-scale initiative as a shared, collaborative programme between EU and participating countries (ELE Programme).**
  - Main scientific goal: *Deep Natural Language Understanding by 2030*.
  - Focus on language equality and the provisioning of technologies, services and resources outside the often-preferred languages. Leverage multi-modal and multilingual resources to support the development of applications for languages and varieties with scarce resources.
- **MT:** We need better *context modeling.* How do *visual features* improve MT? *MT quality:* more focus upon research that looks into the root causes of MT performance issues.
- **Speech:** Development of *natural synthetic voices;* ASR, MT and text-to-speech should focus on a wide variety of languages, including variants and dialects, to keep those alive.
- **Text Analytics:** Research on *DNLU which integrates speech, NLP and contextual information* and *additional modalities*; support basic research in *neurosymbolic approaches*, including *grounding*.

SRIA: Section 6.2 – Research Recommendations 2/3 *(draft)*

- **Create large open access language models for all European languages** (for fine-tuning and downstream tasks): *datasets* (for training and testing), *multilingual models*, models that include *symbolic knowledge*, models that include *discourse features*.
- Combine long-term encyclopedic knowledge with short-term, learned knowledge representations.
- More interdisciplinary research: enable better modeling of multimodal environments, research how modalities can enrich one another and how training and test sets can be constructed.
- Deep learning and neural approaches: focus on trustworthy, interoperable, explainable LT/NLP/AI.
- Bring together the statistical AI and symbolic AI communities to work together on future Semantic AI approaches, especially in multilingual and multimodal environments.
- Combine interactive LT (conversational AI) with text, knowledge and multimedia technologies for a new generation of applications.

SRIA: Section 6.2 – Research Recommendations 3/3 *(draft)*

- Focus on few-shot and transfer learning for settings, in which training data is not available.
- Design a general enough representation of the world around us; build or assemble and convert existing world-knowledge databases to some instances of the common representation; design methods of inferencing and constant updating of such a knowledge base.
- Need for language and culture-specific technologies (not just transfer from English).
- LT should be inclusive and non-discriminatory, it should address safety and security.
- Raise awareness of digital language equality in international LT fora and incentivise Open Access journals and conferences dedicated to less supported and less researched languages.
- Development of technologies covering more linguistic phenomena and text types.
- More systematic inclusion of sign languages.

SRIA: Section 6.3 – Technology and Data Recommendations *(draft)*

- Reduce the technology and resource gap between Europe's languages.
- **Address the *huge* problem of a lack of available data** (relevant for almost *all* European languages).
  - More focus upon **systematic language data collection** (text, dialog, multimodal). Exploit automatic data generation (synthetic data), crowd-sourcing, translation of data.
  - Unleash the power of public sector data, data from broadcasters, social media, publishers etc.
  - Develop methods to overcome the unequal data availability, by focusing on, e.g., annotation transfer, multilingual models preserving quality, few-shot or zero-shot learning.
- Focus upon open ecosystems, open source, open access, open standards, interoperability; foster international standardisation of European approaches in LT, NLP, AI, e.g., APIs etc.
- More systematic research in data bias including dimensions of bias and how to tackle bias.
- Focus upon green LT with a small compute and carbon footprint (e.g., model compression).
- Development of new methodologies for transfer and adaptation of resources and technologies to other domains and languages.

SRIA: Section 6.4 – Infrastructure Recommendations *(draft)*

- **Long-term support of infrastructure initiatives – allows implementation of interoperability.**
  - LT infrastructures & data spaces that support research and development activities (such as ELG).
- **Strengthen and reinforce investment into ELG** as the primary European LT platform; enable and foster the sharing of LT resources, services, datasets, models and code between all stakeholders.
- Infrastructures are essential for maximising the exploitation potential of publicly funded research results. Sufficient operational capacity needs to be ensured, especially for language models.
- Infrastructures should provide interoperability by following standards for data interoperability as well as services and innovative metadata and data management tools for the whole data life cycle.
- **Ensure access to HPC and suitable compute infrastructure** (current HPC centres come with too much administrative overhead and do not meet the needs of ML and LT/NLP researchers).
- Deposit all publicly funded resources in one central LT repository.

SRIA: Section 6.5 – Policy Recommendations *(draft)*

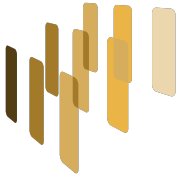
- **Establish a large-scale, long-term coordinated funding programme for research, development and innovation in the field of LT at European, national and regional levels, tailored specifically to Europe's needs and demands to tackle Deep NLU as the next scientific breakthrough.**
  - Increase efficiency by sharing infrastructures, resources and knowledge, with a view to developing innovative technologies and services.
  - Increase (interdisciplinary) collaboration between research centres, academia, enterprises (particularly SMEs and start-ups) and other relevant stakeholders.
  - Boost long-term basic research as well as knowledge and technology transfer between countries and regions.
  - *Efficient and targeted creation of data sets and language resources needs to be supported.*
- Include LT/NLP in all (inter)national AI, language and data policies as well as strategies.

SRIA: Section 6.6 – Governance Model *(draft)*

- **ELE Programme as a shared, collaborative programme between EU and participating countries.**
  - **Level of EU:** Coordination of the ELE Programme, shared platforms and infrastructures (e.g., ELG), shared scientific goals, shared tasks, shared data provision etc.
  - **Level of participating countries:** national and regional funding to develop technologies and resources for all relevant languages of said country; create national centres of excellence.
- Allocate multilingualism and LT to the portfolio of a Commissioner so that LT can get more visibility and recognition in the EU; create a centre for linguistic diversity.
- Require all EU projects to have a language diversity plan so that they show how they will ensure that their results are applicable to people who speak languages other than English. At least three languages to be supported in UI (if any) with one of them from a less-widely spoken language.
- EU could require Member States to procure LT from their local LT industry.
- Learn from DARPA: The decisive factor is the *competitive* scheme. Fund several competing consortia and organise regular technology competitions on especially prepared benchmarks.

European Language Grid

European Language Equality



EUROPEAN LANGUAGE GRID

EUROPEAN LANGUAGE EQUALITY

Thank you!

The European Language Grid has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825627 (ELG).
The European Language Equality project has received funding from the European Union under grant agreement № LC-01641480 – 101018166 (ELE),

Jan Hajič (Charles University), Georg Rehm (DFKI)
hajic@ufal.mff.cuni.cz, georg.rehm@dfki.de

08/09-06-2022 META-FORUM 2022 – Joining the European Language Grid (hybrid conference)
http://www.european-language-grid.eu – http://european-language-equality.eu

Session 10: Language Technology Vision for 2030

14:00 Our Vision for 2030: Digital Language Equality – Jan Hajič (Charles University, Czech Republic)

14:10 Strategic Research and Innovation Agenda – Georg Rehm (DFKI, Germany)

14:30 Discussion

15:00 End of Session 10