DisAI.eu

D4.6
Report on CLARIN integration and other liaison activities

Project Title	Improving scientific excellence and creativity in combating disinformation with artificial intelligence and language technologies
Contract N°.	101079164
Type of Action	HORIZON-CSA
Topic	Disinformation Combating
Project start date	1st Dec 2022
Duration	36 months

Funded by
the European Union

Deliverable title	Report on CLARIN integration and other liaison activities
Deliverable number	D4.6
Deliverable version	1.0
Contractual date of delivery	30 November 2025
Actual date of delivery	28 November 2025
Nature of deliverable	Report
Dissemination level	Public
Work Package	WP4
Task(s)	T4.5
Partner responsible	DFKI
Author(s)	Simon Ostermann (DFKI), Andrej Ridzik (KInIT), Michal Kompan (DFKI), Marian Simko (KInIT)

Abstract	The report outlines the results of the implementation of the measures aimed at improving the visibility and integration of KInIT in major European LT initiatives and bodies, which include LDS, CERTAIN, BDVA, and CLARIN.

Keywords	AI, NLP, CLARIN, LDS, BDVA, CERTAIN, integration

Table of Contents

1 Introduction

This deliverable reports on the activities conducted within Task 4.5, which aims to strengthen the visibility, integration, and strategic positioning of KInIT within major European ecosystems for language technologies (LT) and artificial intelligence (AI). In line with DisAI's objective to advance scientific excellence and support innovative approaches to combating disinformation, the task develops and maintains connections with several key European initiatives. The focus during the reporting period has been on the Common European Language Data Space (LDS), the Centre for European Research in Trusted AI (CERTAIN), the Common Language Resources and Technology Infrastructure (CLARIN), and the Big Data Value Association (BDVA).

At the beginning of the project, CLARIN was expected to be one of the central platforms for KInIT's integration efforts. Its established role in long-term preservation and accessibility of language resources made it a natural point of reference. In practice, however, developments during the implementation phase required a shift of emphasis. For most of the reporting period, Slovakia had not yet joined CLARIN ERIC, which limited the possibilities for formal participation. This barrier was only resolved toward the end of 2025, when Slovakia became a full member and preparations began for establishing the national CLARIN-SK consortium, where KInIT is now involved. At the same time, the broader European environment for data and AI evolved quickly. The launch of the LDS under the Digital Europe Programme and the formation of the CERTAIN network provided new and timely opportunities that aligned closely with the goals of DisAI. These emerging initiatives create more direct pathways to influence data governance, strengthen multilingual AI development, and cooperate with industrial partners, and therefore extend the scope of engagement beyond what CLARIN currently offers.

For these reasons, KInIT concentrated its efforts during this period on deepening its involvement in LDS, CERTAIN and BDVA, while at the same time preparing for future responsibilities within the newly forming CLARIN-SK consortium. This approach ensures that KInIT's expertise in natural language processing, disinformation research and trustworthy AI is applied where it can contribute most effectively.

The remainder of this report outlines the activities undertaken, the progress achieved, and the planned next steps concerning each of the four initiatives. By presenting these developments, the deliverable illustrates how KInIT is positioning itself within Europe's evolving AI and LT infrastructures and supporting both the objectives of DisAI and KInIT's long-term commitment to a multilingual and innovation-oriented European digital ecosystem.

2 Integration of KInIT into the Common European Language Data Space (LDS)

2.1 An Overview of LDS

The **Common European Language Data Space (LDS)** is one of the official Common European Data Spaces launched under the Digital Europe Programme as part of the European Commission's strategy to create a single market for data. It is designed to provide a trusted and interoperable environment for the sharing, exchange, and re-use of language data, models, and services across Europe. The initiative builds upon the achievements of previous European projects such as the **European Language Grid (ELG)** and **European Language Resource Coordination (ELRC)**, extending their scope by introducing a structured governance model, a legal and technical framework for data licensing, and a sustainable operational infrastructure for language technology (LT) stakeholders.

The LDS aims to foster innovation and competitiveness by enabling both public and private organisations to share and access multilingual data and tools while ensuring compliance with European values, privacy, and data protection regulations. It contributes directly to Europe's digital sovereignty and supports the development of multilingual AI systems that can serve the continent's linguistic and cultural diversity. Through its federated architecture, the LDS connects national and sectoral data platforms, ensuring interoperability and promoting collaboration across countries and industries.

The project launched its beta platform in 2024, offering an integrated marketplace where datasets, tools, and services can be discovered, licensed, and re-used. This environment serves as a cornerstone for the future European Language Data ecosystem, linking research institutions, companies, and public administrations that develop or rely on language technologies. By participating in the LDS, KInIT strengthens its position within the European LT landscape, contributing its expertise in natural language processing (NLP), while gaining access to a vibrant network of European partners. This participation enhances KInIT's visibility and supports its mission to help build a more connected, multilingual, and innovative European AI ecosystem.

2.2 Connecting KInIT and LDS

Joining the LDS Interest Group

In July 2025, KInIT became a member of the LDS Interest Group. The **LDS Interest Group** (also previously referenced as "User Group") brings together stakeholders from industry, academia, and public administrations who are actively engaged with the European Language Data Space. Its members help shape and validate the evolving infrastructure by participating in test phases, providing feedback on new features, contributing requirements, and validating use cases. The group plays a bridging role: it ensures that the

LDS platform development is closely aligned with the real needs of data providers and consumers. For instance, its members were invited to earlier rounds of connector testing and discussions on data licensing, metadata standards, and interoperability. As of September 2025, the LDS announced its fourth meeting of the User / Interest Group, indicating that the group continues to be a core forum for stakeholder engagement and platform updates.

Through its participation in the **LDS Interest Group**, KInIT actively contributes to the co-design and refinement of the European Language Data Space. Membership in this group allows KInIT to engage directly with policymakers, technology providers, and other research institutions shaping the governance, interoperability, and operational models of the LDS. By sharing feedback on connector deployment, metadata standards, and data-sharing practices, KInIT helps ensure that the platform meets the real needs of research and innovation actors. At the same time, participation enhances KInIT's visibility and integration within the European LT community, providing early access to platform developments, networking opportunities with leading stakeholders, and a stronger voice in defining the future direction of multilingual and trustworthy AI in Europe

Installing the LDS connector

In October 2025, KInIT installed an LDS connector to connect to LDS and share datasets. The **LDS connector** is the essential software bridge that allows an organisation's local infrastructure to integrate into the European Language Data Space. Once installed on a data provider's or consumer's system, the connector enables the full set of operations defined by LDS, including publishing data descriptions, discovering resources in the marketplace, negotiating licences, and executing secure transfers of language assets. Because the LDS architecture is federated, each participant retains control over their data while interoperating with the broader network via these connectors.

At KInIT, the deployment of the LDS connector represents a key step toward deeper integration into the European language technology ecosystem. It allows KInIT to expose its language resources, tools, and research outputs to the wider community while simultaneously gaining access to datasets and services from other European partners. This participation enhances the visibility and reusability of KInIT's work and contributes to the broader objective of fostering collaboration and innovation within the multilingual AI and language technology landscape in Europe. To date, KInIT has released two datasets via the LDS connector: AMC-16K[1] and a dataset as part of an EMNLP submission.[2]

The Slovak LDS Workshop

As part of its stakeholder engagement strategy, the Common European Language Data Space (LDS) organises a series of Country Workshops across EU Member States and associated countries. These events serve as national entry points to the LDS initiative, aiming to strengthen cooperation with local language technology (LT) communities, public

[1] https://language-data-space.eu/catalogue/view-offer/8692a470-f584-480c-83d4-9fab84a720fe/
[2] https://language-data-space.eu/catalogue/view-offer/8692a470-f584-480c-83d4-9fab84a720fe/

administrations, and industry stakeholders. The workshops provide a forum to present the objectives and architecture of the LDS, demonstrate the functionality of the LDS connector, and discuss opportunities for participation and contribution at the national level. They also facilitate two-way communication between the European LDS consortium and local actors, helping to identify country-specific priorities, challenges, and best practices related to language data sharing and multilingual AI development.

The Slovak LDS Country Workshop took place in Bratislava on 7 November 2024, hosted by the Slovak National Corpus and supported by the European LDS project team. The event gathered representatives from research institutions, universities, public bodies, and private companies active in the field of language technologies. The workshop featured presentations on the overall LDS concept, the Digital Europe Programme, and the emerging European data infrastructure for language resources and services. From the LDS consortium, Simon Ostermann (DFKI) presented the topic *Developing a Market for Language Data and Services*, illustrating how the LDS will enable structured exchange, licensing, and monetisation of language assets across Europe.

KInIT played an active role in the Slovak workshop, both as a presenter and as a national stakeholder, contributing to the discussion on language data and AI. Marián Šimko (KInIT) introduced the DisAI project as an example of an initiative closely aligned with the goals of the LDS, namely, promoting multilingual, trustworthy, and socially responsible AI research. KInIT also participated in a panel session on *Language Data and Language Technologies in Slovakia and for Slovak*, which brought together key national stakeholders to discuss the state and needs of the Slovak LT ecosystem. Through this participation, KInIT contributed its expertise in natural language processing, data ethics, and research management, while reinforcing its role as a bridge between the Slovak and European language technology communities.

KInIT's involvement in the Slovak LDS Country Workshop significantly enhanced its visibility both nationally and at the European level. It positioned the institute as an active contributor to the development of the LDS and as a trusted partner in shaping the future of language data governance and innovation in Europe. KInIT demonstrated its commitment to supporting the goals of the LDS: Building a federated, multilingual data ecosystem that underpins Europe's digital sovereignty and AI competitiveness.

2.3 KInIT and LDS - Next Steps

Following its successful integration into the Common European Language Data Space (LDS), participation in the LDS Interest Group, and active contribution to the Slovak LDS Country Workshop, KInIT is now entering a new phase of engagement focused on data contribution, ecosystem collaboration, and long-term strategic participation.

In the **short term**, KInIT's priority is to expand and refine the publication of its language resources through the operational LDS connector. With the connector now fully installed and functional, KInIT is already releasing its first datasets and metadata into the LDS

ecosystem. The next steps involve ensuring that these resources are discoverable, well-documented, and compliant with LDS metadata and licensing standards. KInIT can also monitor usage and feedback to improve the visibility and usability of its datasets within the platform. In parallel, it can continue to participate in LDS technical and thematic working groups, contributing insights from its practical deployment experience to support the optimisation of the connector and platform functionalities.

In the **medium term**, KInIT can leverage its operational role to become a reference organisation for Slovak language resources within the LDS. This includes supporting other Slovak data holders, such as universities, public institutions, and companies, in connecting to the platform and adopting best practices for data sharing. By facilitating this coordination, KInIT can help establish a stronger Slovak national node within the European LDS ecosystem and ensure that the Slovak language and datasets are adequately represented in multilingual AI development.

In the **longer term**, KInIT can explore opportunities to co-lead thematic initiatives or new projects linked to the LDS, particularly in areas of multilingual NLP, data governance, and ethical AI. Its experience with both data sharing and trustworthy AI research uniquely positions KInIT to bridge the technical and ethical dimensions of language data management. Sustained engagement in the LDS Interest Group, participation in future country workshops, and collaboration with European partners will reinforce KInIT's role as a key contributor to the development of a federated, multilingual, and trustworthy European data ecosystem.

3 Integration of KInIT into the Centre for European Research in Trustworthy AI (CERTAIN)

3.1 An Overview of CERTAIN

The **Centre for European Research in Trusted AI (CERTAIN)** is a European initiative dedicated to advancing excellence and collaboration in trustworthy and reliable artificial intelligence. Established with support from local ministries at Saarland and coordinated by the German Research Centre for Artificial Intelligence (DFKI), CERTAIN brings together leading research institutions, industrial partners, and standardisation bodies to address the growing need for transparent, secure, and certifiable AI systems. Its work aligns closely with European policy frameworks such as the EU AI Act and the Coordinated Plan on AI, focusing on strengthening Europe's ability to design, validate, and deploy AI technologies that are trustworthy by design.

CERTAIN's mission is to bridge the gap between cutting-edge research and real-world deployment by creating methods, tools, and frameworks that allow measurable trust in AI. This includes the development of technical approaches for robustness, explainability, and fairness, as well as engagement with regulatory and standardisation processes to support certification mechanisms for AI systems. The initiative also serves as a platform for cooperation among European AI excellence centres, promoting knowledge exchange and cross-sector collaboration.

Launched officially in September 2023, CERTAIN positions itself as a European hub for trustworthy AI, combining research, policy dialogue, and ecosystem development. Mid 2025, the CERTAIN network was officially kicked off. By joining the CERTAIN network in October 2025, KInIT became part of a pan-European network dedicated to advancing reliable and ethically grounded AI. Its involvement enhances collaboration opportunities with key players in the European AI ecosystem and supports its broader mission of promoting responsible and socially beneficial AI research. Through CERTAIN, KInIT contributes to shaping Europe's collective capacity to build AI technologies that are innovative and aligned with fundamental rights, ethical standards, and public trust.

3.2 Connecting KInIT and CERTAIN

KInIT is an active contributor within the CERTAIN network, aligning its research and institutional strengths with CERTAIN's goal of advancing trustworthy AI across Europe. As a partner in the DisAI project, KInIT is listed among CERTAIN's projects, thereby integrating its work on disinformation, multilingual NLP, and ethical AI into the broader agenda of reliable AI systems. Through this affiliation, KInIT gains access to shared resources, collaborative research opportunities, and participation in working groups focused on robustness, interpretability, human oversight, and hybrid AI techniques.

Moreover, KInIT's network presence within CERTAIN enables it to interact with other leading research institutions, industry partners, and standardisation entities that contribute to the certification and formalisation of trustworthy AI. CERTAIN's mission includes both fundamental research and also "defragmentation" of the European AI trust landscape and promoting collaboration among actors with diverse perspectives. By positioning itself within this structure, KInIT can both inform and benefit from the alignment of research, regulation, and applied deployment in trustworthy AI across Europe.

3.3 KInIT and CERTAIN - Next Steps

Building on its active involvement in the Centre for European Research in Trusted AI (CERTAIN) network, KInIT is well positioned to expand its role in advancing research and collaboration on trustworthy and responsible AI in Europe. CERTAIN's mission—to foster excellence, interoperability, and impact in the domain of reliable and certifiable AI—closely aligns with KInIT's institutional focus on AI ethics, transparency, and explainability, making this partnership strategically important for both scientific and policy-oriented engagement.

In the **short term**, KInIT can deepen its engagement by contributing to CERTAIN's working groups and thematic clusters, particularly those addressing explainability, robustness, and human oversight in AI systems. By sharing its expertise from projects such as DisAI and its practical experience with ethical governance frameworks, KInIT can help ensure that the CERTAIN network integrates both technical and societal dimensions of trustworthy AI. Regular participation in consortium meetings, workshops, and policy dialogues will strengthen KInIT's visibility and influence within the community.

In the **medium term**, KInIT can aim to co-develop research prototypes or pilot activities under CERTAIN's umbrella that demonstrate concrete applications of trustworthy AI methods—especially in the context of combating disinformation, multilingual language technologies, and human-centric AI design. These pilots would not only validate scientific approaches but also provide practical evidence supporting future certification schemes and EU policy initiatives related to the AI Act.

In the **longer term**, KInIT can work toward becoming a regional competence hub for trustworthy AI, connecting Central and Eastern European actors to the wider CERTAIN network. By helping to disseminate best practices, hosting joint training or summer schools, and contributing to white papers and position statements, KInIT can consolidate its leadership in responsible AI research and advocacy. Sustained participation in CERTAIN will thus allow KInIT to shape the future European landscape for safe, transparent, and accountable AI, while reinforcing its visibility as a key contributor to the EU's trustworthy AI ecosystem.

4 Integration of KInIT into the Common Language Resources and Technology Infrastructure (CLARIN)

4.1 An Overview of CLARIN ERIC

CLARIN (Common Language Resources and Technology Infrastructure) ERIC (European Research Infrastructure Consortium) is a European research infrastructure that links together centres such as universities, libraries, archives, and research institutes to provide long-term access to digital language data and language technologies. Through this federation of centres, researchers can work with text, audio, video and other language resources using compatible tools and services that can be combined across national boundaries.

Formally established as a European Research Infrastructure Consortium in 2012, CLARIN focuses on supporting data-driven research by offering a stable technical environment, persistent storage, and interoperable services. Many of these services are accessible via single sign-on, so users from member countries – and often beyond the academic sector – can discover, analyse, enrich, and reuse language resources regardless of where they are physically hosted.

CLARIN promotes open science and encourages interdisciplinary and comparative research. In cooperation with initiatives such as the European Open Science Cloud and related infrastructures in the social sciences and humanities, it works to ensure that data formats, metadata standards, and research workflows are interoperable and reusable over the long term.

A key role in this ecosystem is played by the technical service centres, known as CLARIN B-centres. These units – typically universities, research institutes, or similar organisations – form the operational backbone of the infrastructure by offering reliable, long-term access to language resources, processing services, and expert support for the research community. To be recognised as a B-centre, an institution must meet strict requirements regarding technical robustness, organisational stability, and quality of services; only centres that pass a formal assessment receive this status.

Currently, there are around two dozen certified B-centres, along with several candidate centres preparing for assessment. Across the network, partners coordinate not only on technical standards, but also on legal, ethical, and training-related aspects of data acquisition and reuse. Joint projects help to develop new resources and tools and to refine best practices.

4.2 Connecting KInIT and CLARIN – Current Status

The membership in CLARIN ERIC is at the country level, not the institutional level. Individual countries organise themselves into national consortia that act as nodes in an international network for the exchange and joint provision of language data and language technologies for research purposes.

In Slovakia, governmental support was not available for a very long time. Repetitive negotiations have been conduced by the Ľudovít Štúr Institute of Linguistics (ĽŠIL), Slovak Academy of Sciences, where Slovak national corpus resides. Its role is to develop, manage, and provide public access to various language resources for Slovak. As an institution of general public interest it has been co-funded by the Ministry of Education, Research, Development and Youth of the Slovak Republic (MŠVVaM SR) and the Ministry of Culture of the Slovak Republic (MK SR).

KInIT has been collaborating with ĽŠIL in the past, while writing common grant proposals or participating in Slovak NLP Community activities since 2024. In February 2025, KInIT expressed preliminary interest in joining a national consortium with ĽŠIL. In September 2025, Slovakia finally joined CLARIN ERIC as its 27th full member[3]. Currently, there is ongoing CLARIN-SK consortium formation, with Ľudovít Štúr Institute of Linguistics as coordinator and KInIT, other institutes of Slovak Academy of Sciences and Technical University of Košice as expected members. The formal agreement was already proposed, and it is expected to be finalised by the end of 2025.

4.3 KInIT and CLARIN – Next Steps

With Slovakia's recent accession to CLARIN ERIC and the ongoing formation of the CLARIN-SK national consortium, KInIT is entering a new phase of engagement in the European language resource infrastructure. As one of the key research institutes in the emerging consortium, KInIT is among the major candidates for hosting the national CLARIN B-Centre—an infrastructure component focused on the long-term preservation, management, and dissemination of language resources. This potential role aligns with KInIT's broader infrastructural ambitions in data stewardship and its commitment to strengthening the Slovak NLP ecosystem.

In the **short term**, KInIT can deepen its involvement by actively contributing to the design of the CLARIN-SK governance and technical model, coordinating closely with ĽŠIL and other partners. This includes assessing the technical, organisational, and operational requirements of a B-Centre, exploring interoperability with neighbouring CLARIN nodes (especially in the Czech Republic), and identifying how CLARIN's repository standards – such as metadata harmonisation, persistent identifiers, and controlled-access mechanisms – can enhance KInIT's internal data management workflows.

In the **medium term**, KInIT can leverage CLARIN to consolidate and curate its stable language datasets, storing them in a structured, versioned, and well-documented manner that supports future reuse and long-term preservation. While experimental or computationally intensive data work will continue to take place within KInIT's dedicated AI infrastructure, CLARIN can serve as the authoritative repository for final, validated datasets. By demonstrating strong data stewardship practices and ensuring visibility of its resources through European discovery platforms, KInIT can become a central contributor to the newly forming CLARIN-SK infrastructure and help motivate other Slovak organisations to share their own datasets.

[3] https://www.clarin.eu/content/clarin-newsflash-september-2025

In the **longer term**, KInIT can build on this foundation to establish itself as a national competence centre for language resource management and FAIR data practices. This may include supporting Slovak research institutions in onboarding their datasets to CLARIN, hosting training activities, participating in CLARIN's European working groups, and contributing to methodological and technical developments within the network. Sustained engagement will allow KInIT to strengthen its role in shaping Slovakia's language technology infrastructure, improve the visibility and impact of Slovak linguistic resources, and ensure that they are effectively integrated into the wider European research landscape.

5 Integration of KInIT into the Big Data Value Association (BDVA)

5.1 An Overview of BDVA

The Big Data Value Association (BDVA) is an industry-driven non-profit organisation with over 250 members across Europe, including large companies, SMEs, start-ups, research institutions, and user organisations. Its mission is to foster innovation in data and artificial intelligence (AI) by bridging technology, policy, research, and industry. BDVA emphasizes a multidisciplinary approach, combining technical, legal, business, and societal perspectives to address challenges in the data economy. Since its creation in 2014 as the private counterpart to the European Commission in the Big Data Value Public-Private Partnership, BDVA has been instrumental in shaping research agendas, innovation roadmaps, and guidelines for policymakers, while promoting experimentation, collaboration, and co-learning in areas such as big data technologies, industrial AI, and data-driven value creation.

BDVA also acts as a hub for European-level cooperation, offering tools and frameworks to support the co-creation of pan-European data-driven applications and services. Its governance structure includes a General Assembly, Board of Directors, and Task Forces, which channel member input into strategic initiatives. Flagship events like the European Big Data Value Forum (EBDVF) and Data Week provide platforms for knowledge-sharing, collaboration, and visibility at the European level. BDVA is also a founding member of the AI, Data and Robotics Partnership and a private member of the EuroHPC Joint Undertaking, reinforcing its role as a central actor in Europe's digital transformation. Ultimately, BDVA positions itself as the go-to place for shaping the future of European data and AI, with a strong commitment to sustainability, competitiveness, and global impact.

5.2 Connecting KInIT and BDVA

KInIT has applied for the membership in early 2023. It has been successfully approved, and since May 2023 it is an associate member[4]. Thanks to extending the capacity of KInIT in policy advising, advocacy and networking, we foresee BDVA as a unique opportunity for increasing the impact of the Institute.

5.3 KInIT and BDVA – Next Steps

KInIT will benefit significantly from BDVA by gaining access to a strong European network of research institutions, companies, and policymakers focused on data-driven innovation and artificial intelligence. Through BDVA, KInIT will be able to collaborate on EU-funded projects, strengthen its visibility as the Slovak leader in trustworthy AI and data science, and contribute to shaping European standards and policies around data governance and

[4] https://bdva.eu/news/bdva-welcomes-8-new-members/

ethics. Membership also provides opportunities to share expertise, learn best practices from other members, and position KInIT strategically within Europe's evolving AI and data ecosystem, ensuring its research and innovation efforts align with continental priorities and open doors to new funding and collaboration opportunities.

In the **short term** KInIT acknowledged the important role of BDVA in the Institute's mission. With staff capacity increased, we foresee a deeper integration in the BDVA structures in the short term period. First of all, we will explore active participation in the selected "Cross Sectorial Task Forces". Among the ones available, "Data and AI Technologies", "etami", and "Generative AI/Foundation models" are of primary interest.

In the **medium term** KInIT will further strengthen the participation in more task forces, specifically "Policy and Societal", "Data protection", and "Standards and Benchmarking". This participation will, together with other associations in which KInIT participates, further strengthen its position as a hub for trustworthy AI.

In the **long term**, the role of the Institute will also turn towards taking the responsibilities and opportunities in contribution to the organisation structures. In this way, it will become an internationally recognised and relevant player in the R&D area.

5 Conclusions

The integration activities presented in this deliverable reflect a strategic effort to position KInIT within European infrastructures for language technologies and artificial intelligence. The European environment for AI and language technologies evolves quickly, with new mechanisms for data sharing, interoperability, and trustworthy AI taking shape regularly. KInIT adapted to these developments by concentrating on initiatives that offer operational relevance and opportunities for collaboration with both public and private stakeholders.

While **CLARIN** (Common Language Resources and Technology Infrastructure) has long provided an essential research infrastructure for linguistics and digital humanities, its primary focus remains on academic resource curation and long term preservation. In contrast, the **Common European Language Data Space** (LDS) and the **Centre for European Research in Trusted AI** (CERTAIN) represent the next generation of European collaboration frameworks, built to directly support innovation and cross-sector application of AI and language technologies. The LDS extends beyond the traditional research community by creating a federated, market oriented data ecosystem under the Digital Europe Programme, enabling stakeholders to share and monetise language data in compliance with European data standards. It bridges research and industry and aligns with Europe's broader strategy for digital sovereignty and multilingual AI development, which goes beyond the scope of CLARIN's academic infrastructure model. CERTAIN focuses on trustworthy and policy aligned AI, connecting research excellence with emerging European regulation such as the EU AI Act and with industrial adoption of responsible AI technologies.

Given this alignment, KInIT decided to prioritise its institutional engagement on LDS and CERTAIN. Both initiatives provide stronger links between research and innovation, direct pathways for collaboration with industry, and active roles in shaping European AI policies. This strategic focus enables KInIT to contribute where its expertise in AI ethics, NLP, and data driven innovation can have the greatest European impact, supporting the development of open, trustworthy, and application oriented AI ecosystems that complement and extend the foundations established by CLARIN.

Participation in BDVA further broadens this engagement through activities in data driven innovation, policy dialogue, and community coordination. This involvement gives KInIT access to a wider network of industrial and academic partners and strengthens its contribution to European discussions on standards, governance, and responsible AI deployment.

Although KInIT's practical involvement in CLARIN remained limited during most of the reporting period due to Slovakia's late accession to CLARIN ERIC, the formation of the national consortium now provides a basis for future contributions. This development creates a channel for structured, long term integration of Slovak language resources into established European research infrastructures.

Taken together, engagement in LDS, CERTAIN, BDVA, and CLARIN positions KInIT within multiple layers of the European AI and LT landscape. These connections support the objectives of DisAI and reinforce KInIT's role in shaping research and innovation agendas related to data governance, multilingual AI, and trustworthy artificial intelligence across Europe.