AI Survey 2025: analysis

CENL "AI in Libraries" network group

January 2025

---

1. Introduction

This report is the analysis of the CENL network group "AI in Libraries" survey of CENL members regarding their AI practices and maturity. This survey is the second edition of the first launched in September 2021.

Its purpose is based on two objectives:

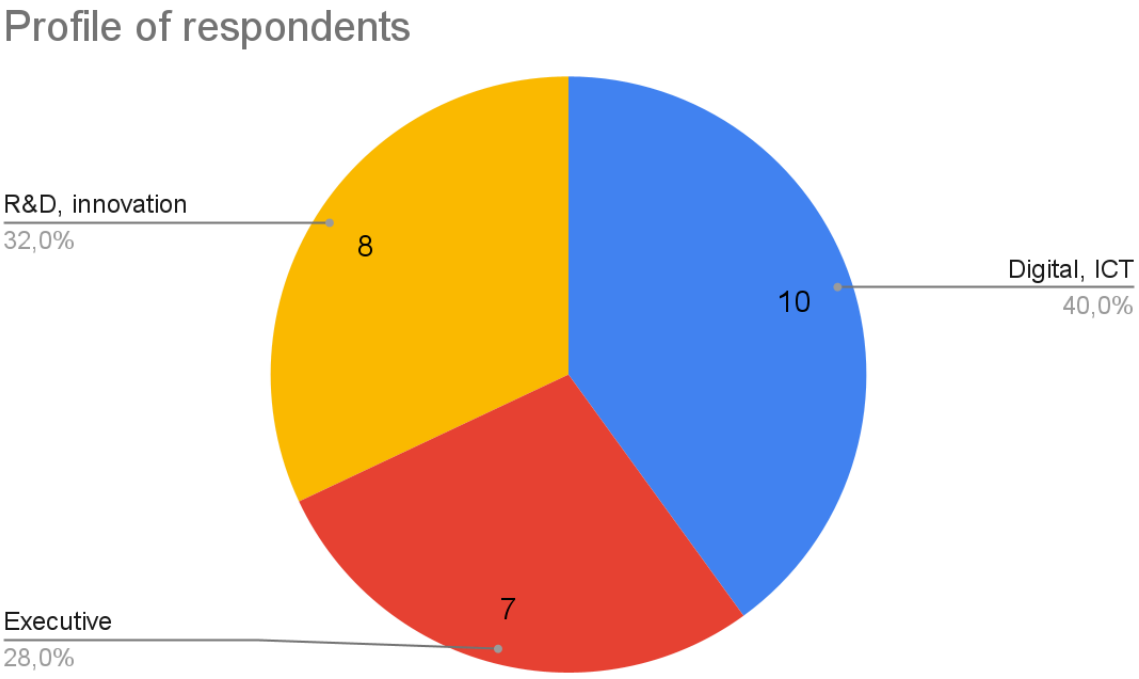
- Identify the current activities or intentions of European national libraries regarding the use of artificial intelligence.
- Make CENL members aware of the existence of this group dedicated to AI.

The report provides an overview of this survey, which received 26 responses from the CENL network of 46 national libraries (NLs) from 45 countries.

2. Profile of respondents

The following graph shows the functions occupied by the respondents to the survey, grouped into three main categories:

- Executive management: director, general director, head of library services…
- Digital, ICT: head of digitization, ICT director, head of data dpt, head of digital collection…
- R&D, innovation: head of digital research, head of datalab, head of digital innovation, research coordinator, AI advisor, data scientist…

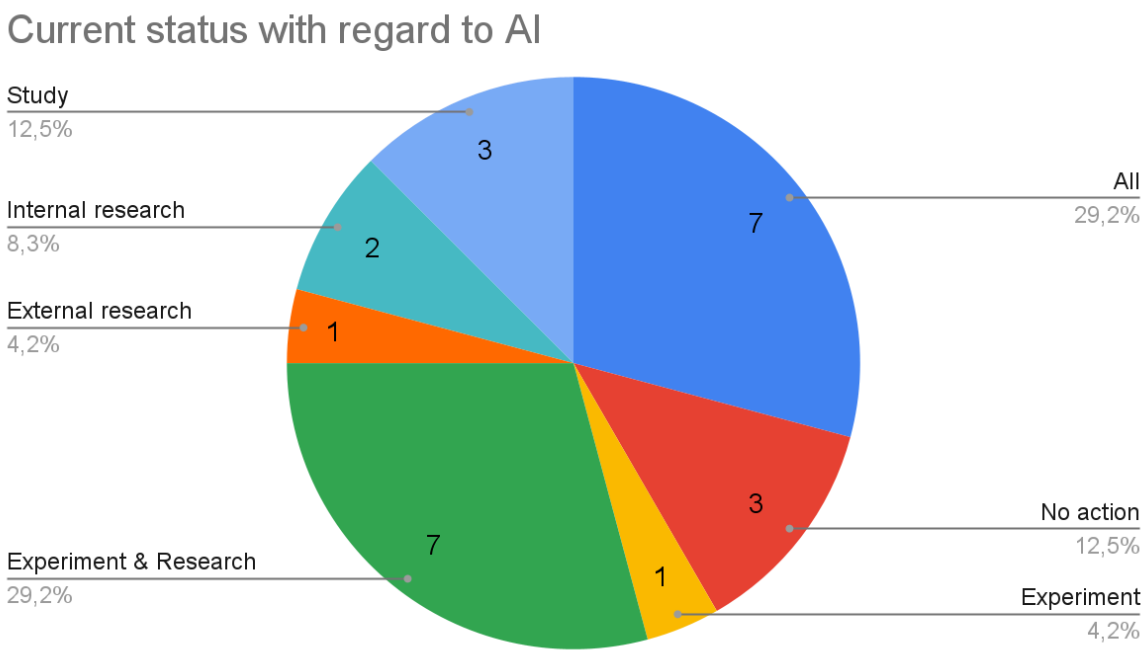


Profiles of respondents (25 answers)

Answers to the survey included responses from the following countries: **Armenia**, Austria, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, **Italy**, **Latvia**, Lithuania, Luxembourg, **Macedonia**, Malta, Netherlands, Norway, Poland, **Romania**, Scotland, Spain, Switzerland, **Ukraine**, United Kingdom (new entries in bold).

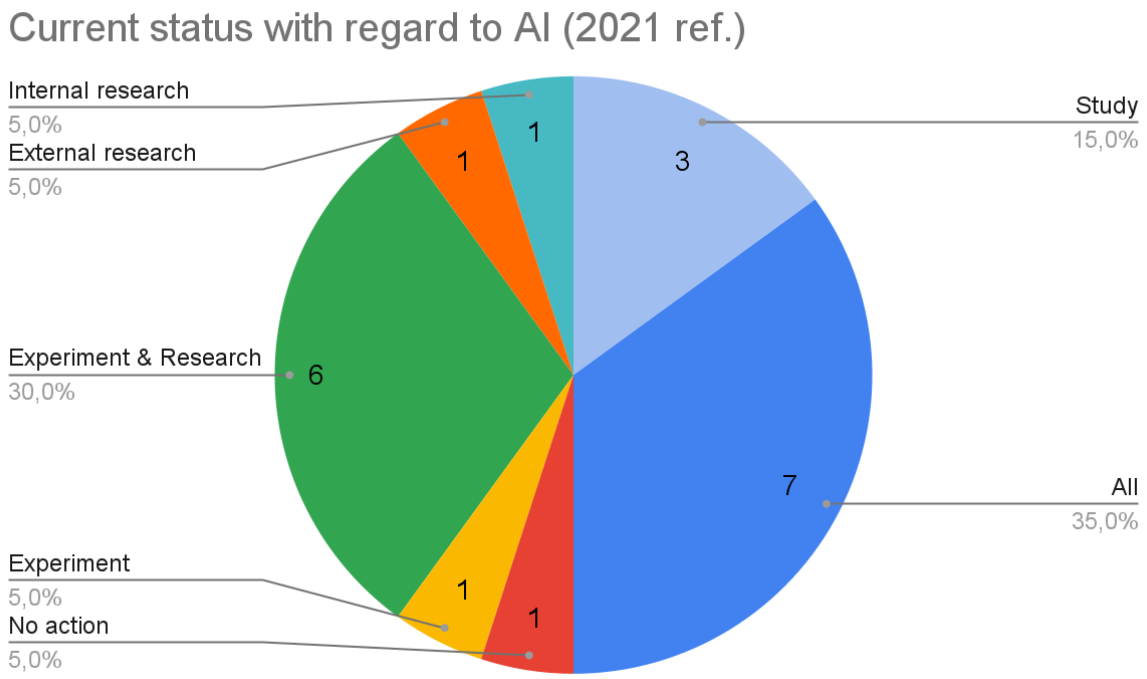
3. Degree to which AI is taken into account in your institution

The raw figures from the survey show that a large majority of libraries have already set up an AI activity (88,5%). Only 3 respondents said they had no AI activity to date. Around 29% cover all types of activity, including AI in production, which is a significant increase compared to the 2021 survey (9%).



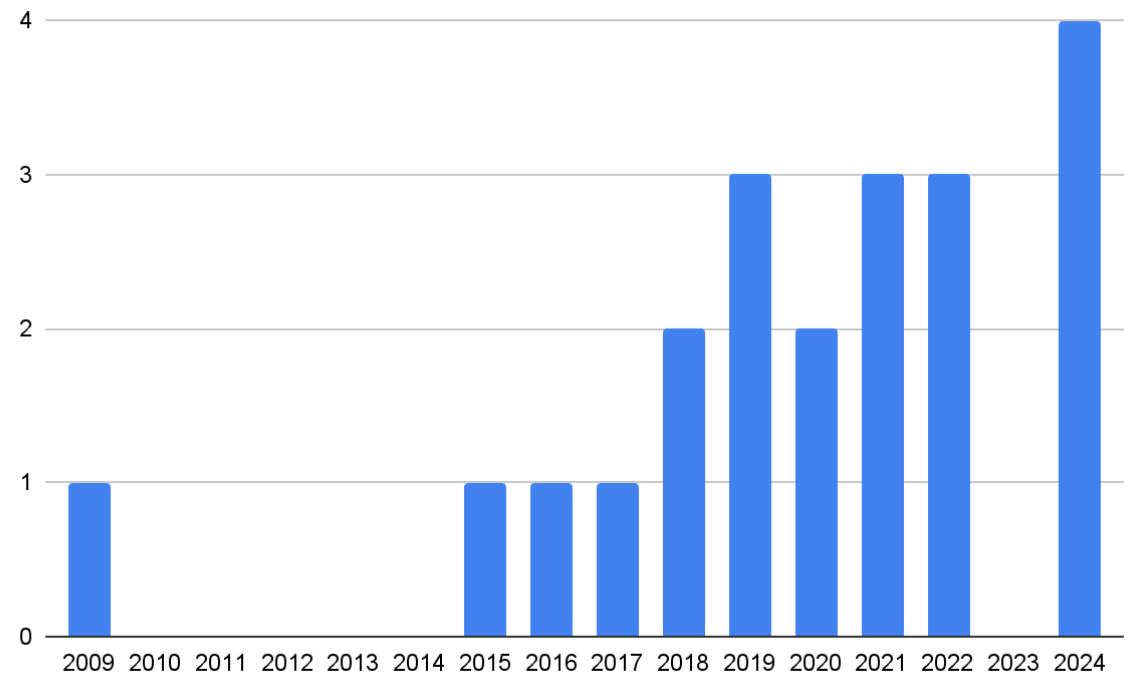
Current AI activity types (multiple answers allowed, 25 answers)

If we try to compare the evolution of the group of 2021 respondents who also took part in this survey, the evolution is even clearer, with ‘no action’ for only 1 member and 35% for all types of activity.



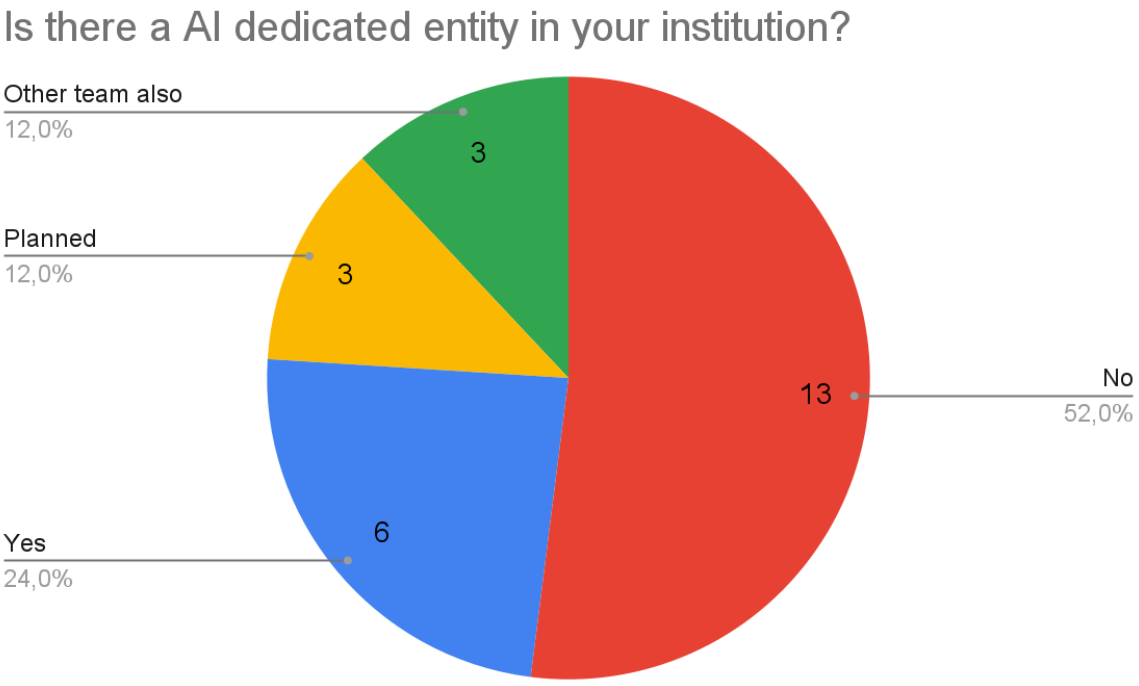
Current AI activity types restricted to the 2021 respondents (20)

The survey data also allows us to assess when NL interest in AI began, i.e. mainly from 2015 onwards:



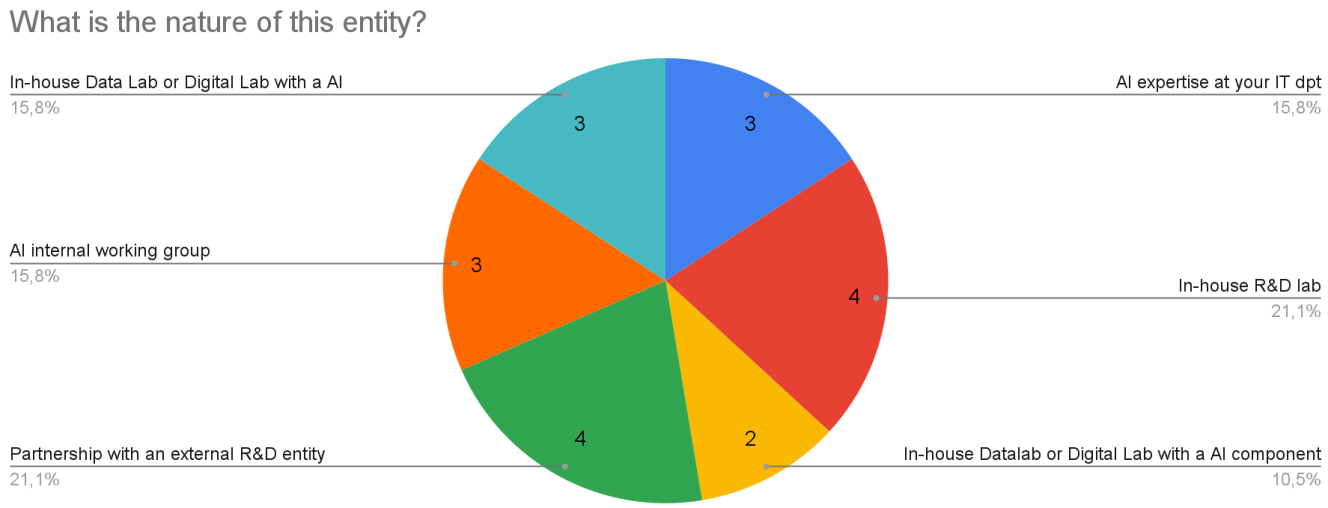
Start date of the AI awareness (21 answers)

To the question of whether NLs do not have an entity dedicated to AI, the responses show that 52% of respondents are in this situation. In 2021, this number was 48%. But if we restrict again the study to the 2021 respondents, it falls to 42%.



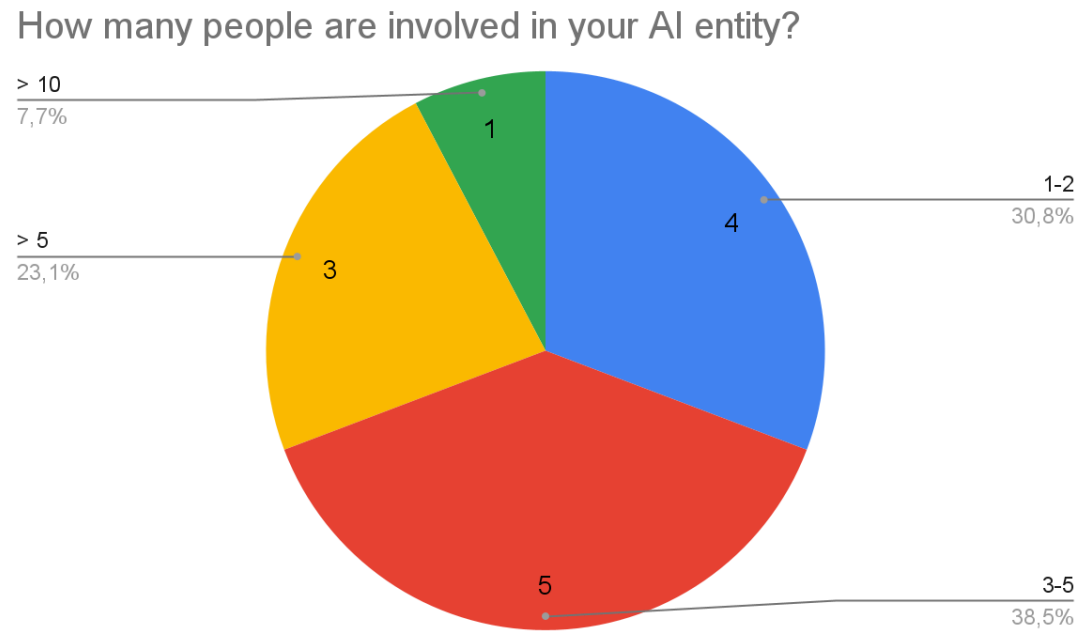
AI dedicated entity (25 answers)

This entity can take different forms, the most common being an internal laboratory or a collaboration with research projects.



Nature of the AI entity (12 answers, multiple answers allowed)

Its size also varies, from 1 to more than 10 people. Most entities (69%) have between 1 and 5 people.



How many people are involved in your AI entity? (13 answers)

Open question: *Add any additional information about your AI maturity that might be helpful (partnership with computer science laboratories, participation in AI initiatives at the national level, training program...) (15 answers)*

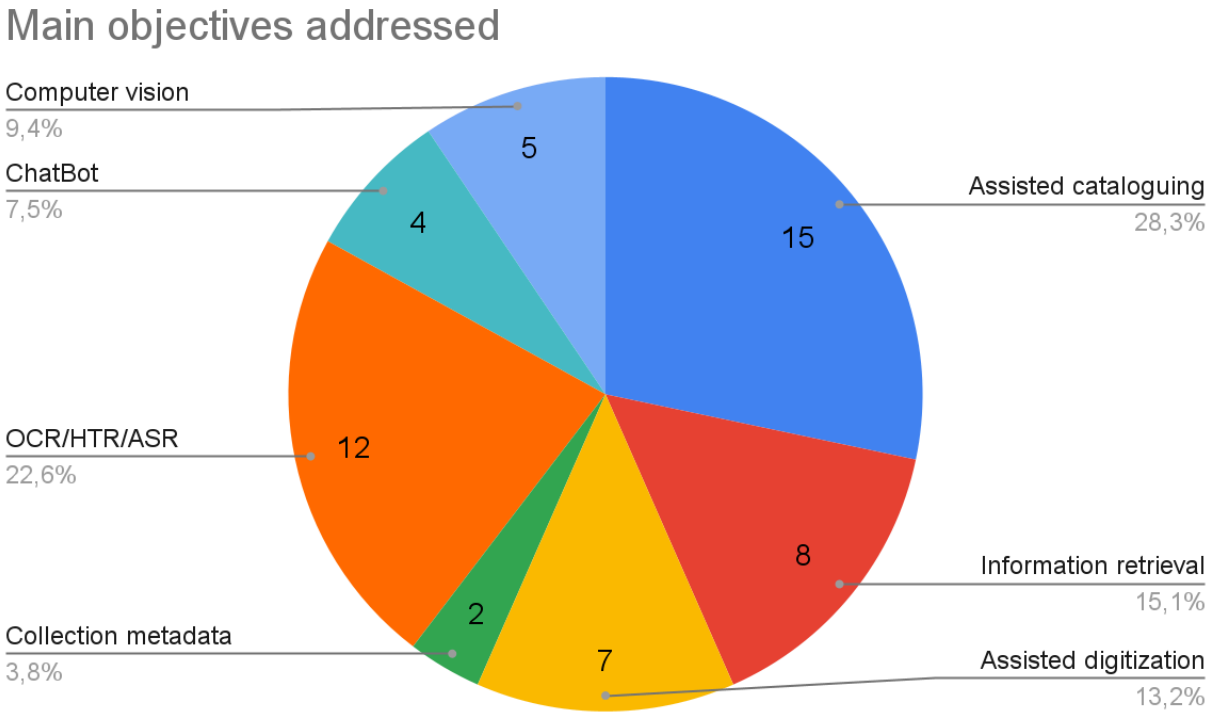
- National Library of Luxembourg: Participation in the national AI4GOV program in 2020-2021
- National and University library "St.Kliment Ohridski" Skopje, Macedonia: More intensive education of employees is planned through cooperation with experts from the academic community, involvement in national and international projects and workshops in the field of AI.
- National Library of Estonia's Digilab: https://digilab.rara.ee/en/
- National Library of Scotland: Hosted AI Symposium (2023), Creation of AI Statement for Library (June 2024), Participation in various research projects involving AI tools (Ongoing since 2019)
- National Library of Czech Republic: we cooperate with the Czech Technical University in Prague
- National Library of Poland: Participation in AI project at international level, contact with computer science companies
- The British Library: Participation in knowledge networks, e.g. AI4LAM since 2018 and development of GLAM specific training resources for AI/ML. Have been involved in research funded projects with partners.
- National Library of Finland: We have developed an automated indexing tool (Annif / Finto AI) which is in production. We have several past and present AI-related projects with partners such as CSC - the IT center for science.
- BnF: we’re implementing our AI roadmap (2021); multiple partnerships; state funded AI projects; internal projects
- Royal Library of Belgium: DSL has been mainly involved in AI from an image processing aspect, especially in collaboration with DIMA/MaDs research group in the math department of Vrije Universiteit Brussel (VUB)
- National and University Library in Zagreb: There is a short online training program named "Understanding and applying artificial intelligence (AI) in the library environment" within the National Center for Continuous Professional Development of Librarians at our Library. The webinar gives the participants the basic theoretical knowledge needed to understand and apply artificial intelligence in their libraries.
- Royal Danish Library: AI and Data initiative is a part of the new 2024-2027 library strategy. First 3 pilot projects running right now. Includes a competence development plan for both the broader scope of employees of the library as well as specific initiatives for e.g. it developers.
- National Library of Norway: We just reached a point in which the unit needs to grow and split into inference/applications and research only tracks
- National library of the Netherlands: We participate in several (research) initiatives and have agreements with at least two companies

- National library of the Spain: We are collaborating with the Secretary of State for Artificial Intelligence and the Ministry of Culture in order to legislate the use of AI and develop a project devoted to the corpus data access

---

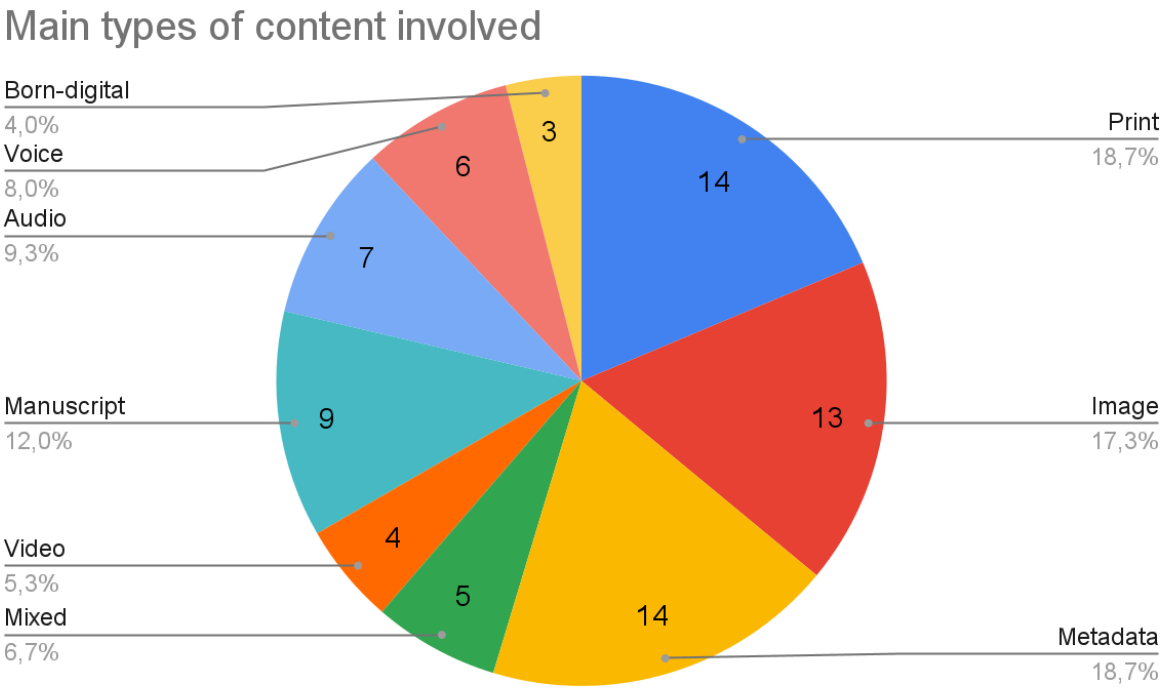
4. AI activity

Our first question on the activity in the NLs is related to AI projects driven by technology (OCR, HTR, etc.). Compared with the 2021 survey, we can see the emergence of new technologies: ASR (automatic speech recognition), chatbot, information retrieval (RAG)



Description of the main current AI activity/projects (24 respondents, multiple answers allowed)

When analyzed in terms of the types of content or collections involved, the ongoing AI activity can be summarized with these numbers:



Types of content involved in the projects (22 answers, multiple answers allowed)

Compared with the 2021 survey, the most targeted types remain the same (print, image, manuscript, metadata) but we are seeing the emergence of the born-digital category (ebooks, PDF, thesis, web archives).

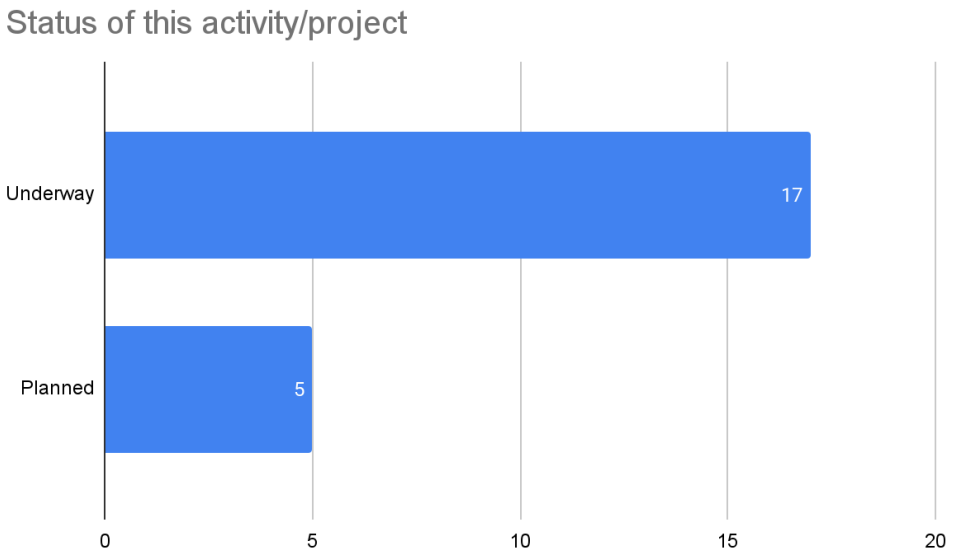
Open question: *Description of one of your main current AI activity/project (depending of your agenda, it can be a study, a project, a partnership…) (22 answers)*

- National Central Library of Florence: Study and prototype of an automatic semantic indexing system for texts in italian language; study on named-entity recognition.
- National Library of Luxembourg: Currently the library is working on two proof of concepts. Autocat is a project that aims to automate as much as possible the cataloguing of the digital legal deposit according to RDA in Marc21. The other one is semantic search in the images in the digitized collection and possibly the addition of the images to the chatbot.
- Swiss National Library: Semi-automatic indexing of digitised posters
- National Library of Estonia: Automated cataloging and subject indexing
- National Library of Scotland: The MapReader app will improve the ability to use machines to automatically identify features on historic maps. First developed in partnership with the Living with Machines project.
- National Library of Czech Republic: Automated subject headings generation from printed documents
- National Library of Poland: A study on implementing AI in cataloguing

- The British Library: FRAIM: Framing Responsible AI Implementation & Management
- Martynas Mazvydas National Library of Lithuania: We are creating project „ e-culture” with over 19 big partner-institutions in Lithuania. It will contain all kinds of digitalized cultural heritage objects of Lithuania.
- National Library of Armenia: Project of Armenian printed texts OCR and voicing
- National Library of Finland: Development of an automated cataloguing tool for subject indexing: We have been developing the Annif tool since 2017 and it has been in production since 2020. Currently we are looking to expand to automatic retrieval of bibliographic metadata using large language models.
- https://www.dnb.de/ki-projekt : Improve automated subject indexing with novel AI methods
- BnF: Large Image analysis project on all the digital library content
- Royal Library of Belgium: Newspaper front page recognition with AI explainability
- Royal Danish Library: Access to the archives of the danish national broadcaster (DR) - potentially a collection of 4 million hours of AV-content that is lagging metadata and have a number of problems regarding the older parts digitized from physical tape carries
- National Library of Norway: ASR for low-resources languages, study on copyright's impact on LLMs
- V. I. Vernadskyi National Library of Ukraine: Improving search capabilities in digitized library collections via AI image recognition apps
- KB, national library of the Netherlands: Finalising a roadmap
- Austrian National Library: Project on object detection and extraction
- National Library of Latvia: Institute of Bibliography is conducting experiments exploring the application of artificial intelligence in cataloguing. (We were inspired by the CENL "AI in Libraries" webinar with Hannes Lowagie (KBR — Bibliothèque Royale de Belgique)

Technologies used in these big projects are very diverse, from proprietary software (GPT, ChatGPT, Microsoft Power Platform) to open source resources (Tensorflow, PyTorch, MLFlow…), community driven products (Annif) or open source models (). The emergence of the large language models (Llama, Mistral, GTP) and multimodal models (CLIP, whisper) in this survey is particularly noteworthy.

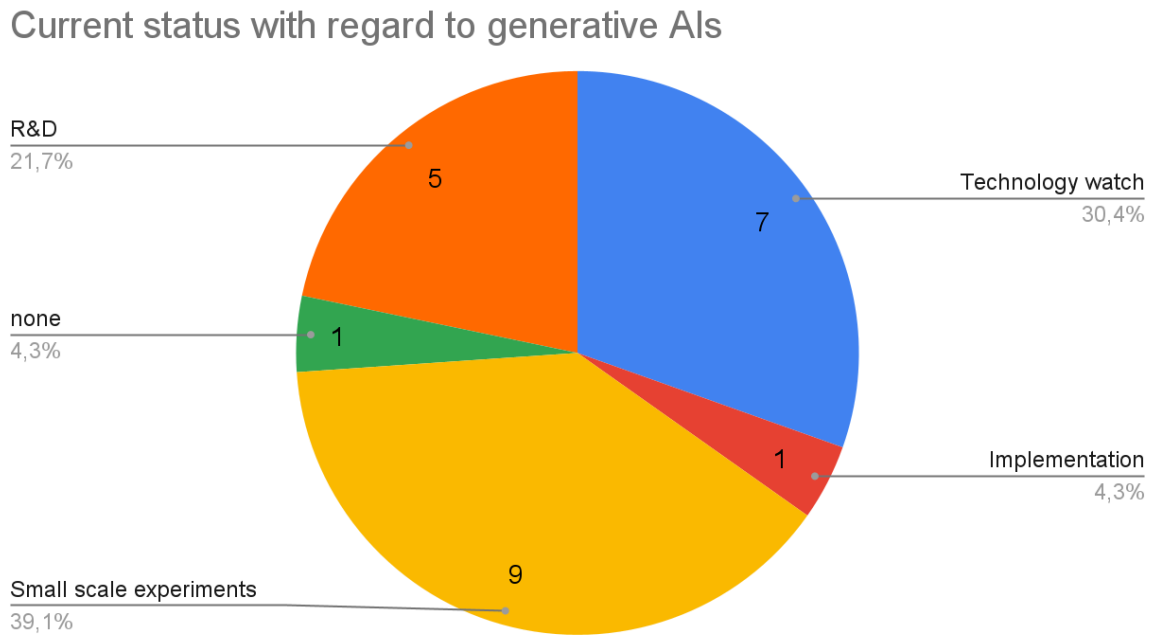
Finally, all these projects or activities are either underway or planned.



Status of this activity/project (22 answers)

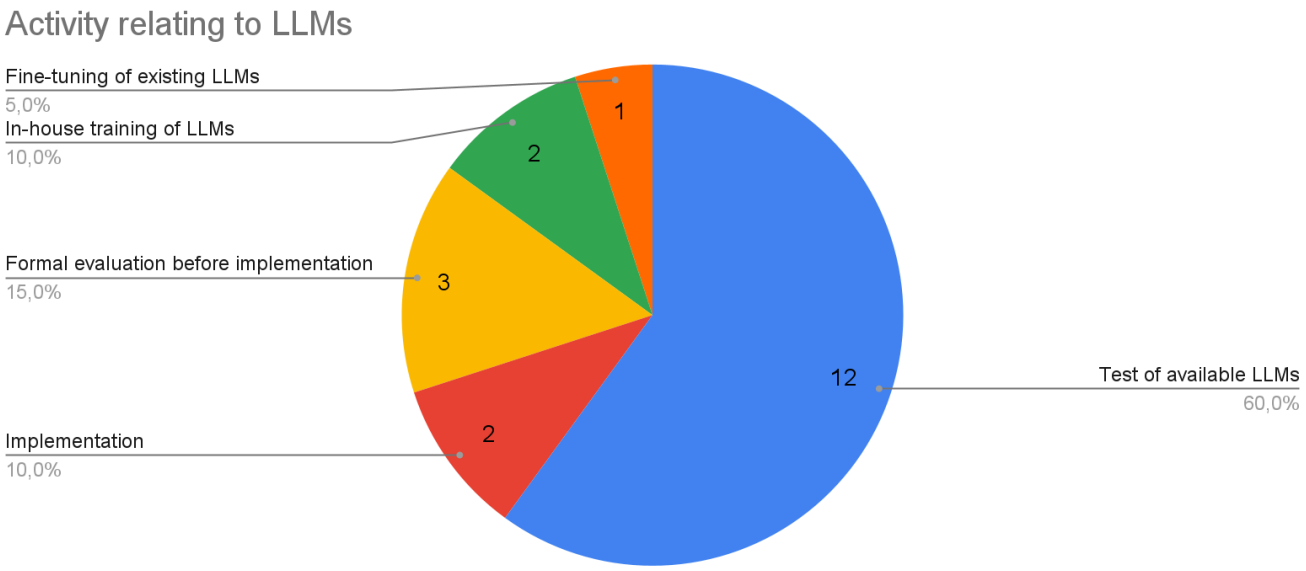
5. Generative AI

The generative AI revolution of recent years has led to a dedicated section in the new edition of this survey. Most respondents have already taken this development into account.



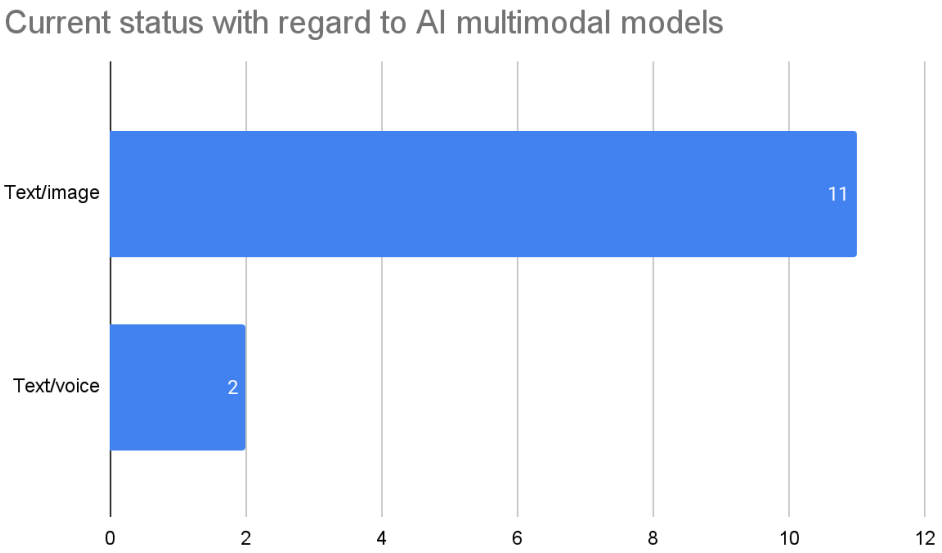
Status with regard to generative AI (23 answers)

In practice, activity around generative AI involves testing existing models, sometimes training, and more rarely going into production.



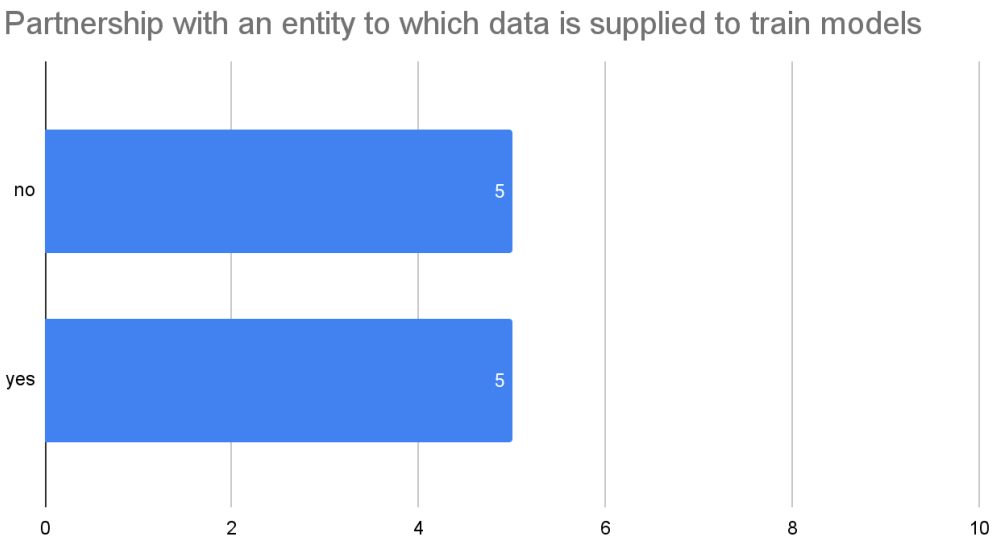
Activity relating to generative AI (22 answers)

Multimodal AI is also addressed by some members (13 positive responses), with text/image models attracting the most interest.



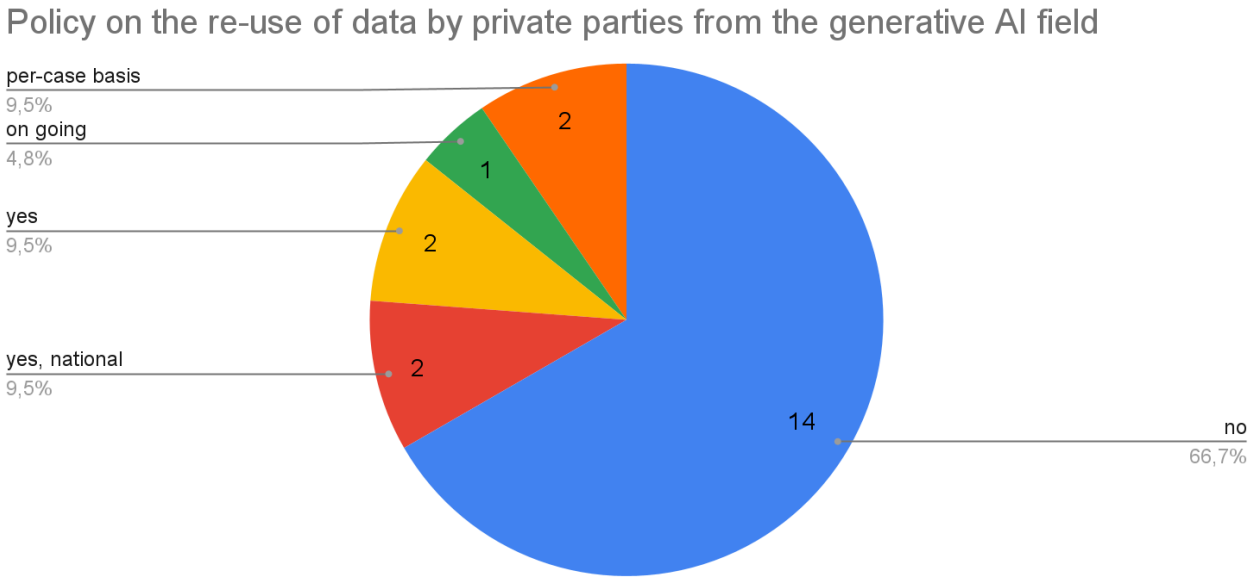
Current status with regard to AI multimodal models (13 answers)

In half the cases, an external entity (laboratory, company) is a partner with which models are produced.



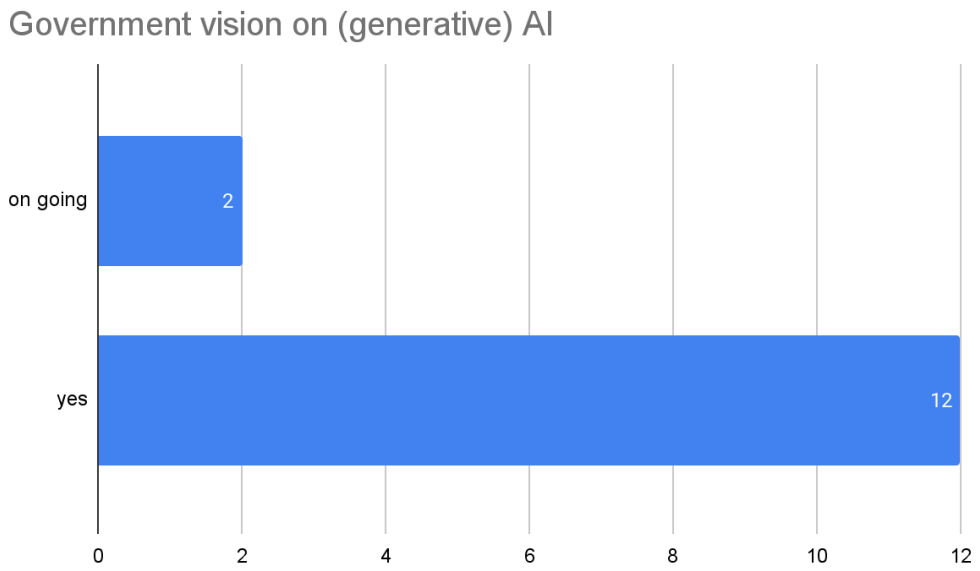
Partnership with an entity to which data is supply to train models (20 answers)

With regard to the policies to be applied to the re-use of content for the AI ecosystem, the landscape of responses is varied, with a predominance of members who have not yet addressed the issue. In other cases, this response is sometimes based on the national level. Very few have made public their own AI policy (1).



Policy on the re-use of data by private parties from the generative AI field (22 answers)

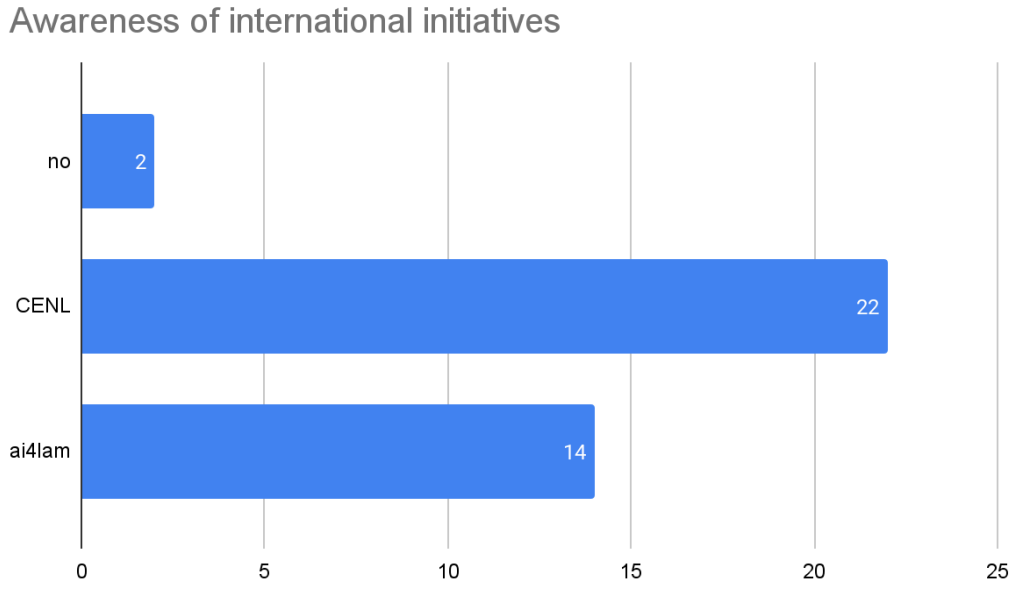
With regard to government policies, 14 members are aware of a current or existing national policy or plan on AI and/or generative AI.



Government vision on (generative) AI (17 answers)

6. Interaction with the CENL “AI in Libraries” network group

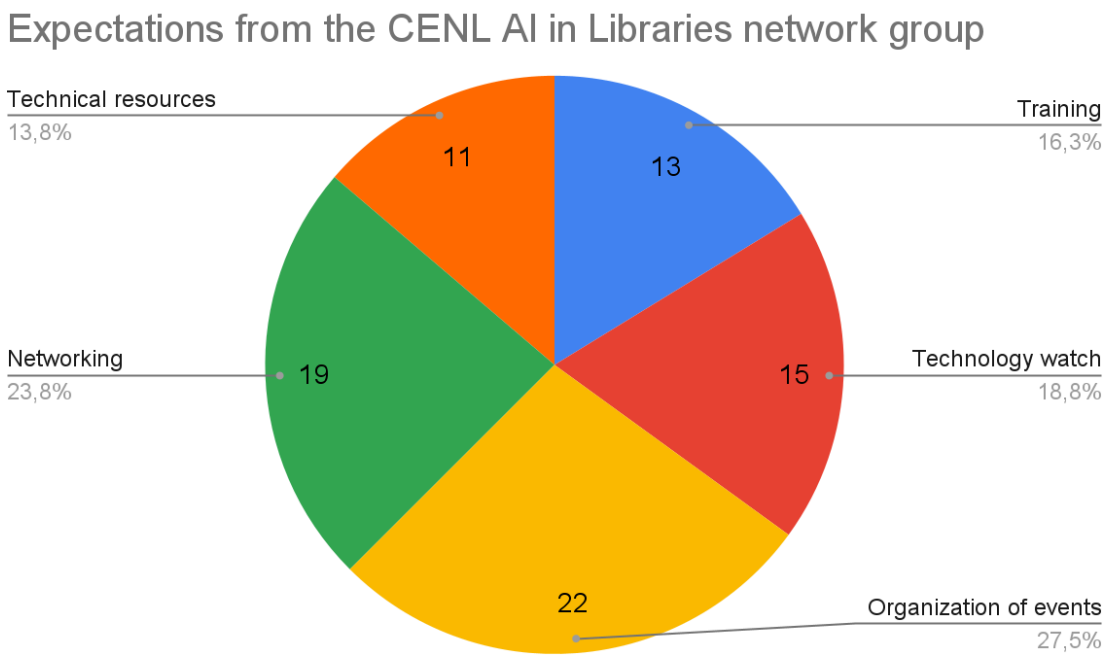
Almost all the respondents were aware of the CENL AI network group (92%). The international AI4LAM initiative is less known (58%).



Awareness of international initiatives (23 answers)

When asked about their expectations and needs with regard to the CENL “AI in Libraries” network group, respondents divided their answers evenly among the actions mentioned in

the survey: building and dissemination of resources, organization of events, networking and sharing.



Expectations from the CENL AI in Libraries network group (25 answers, multiple answers allowed)

Open question: *What could/would you contribute with? (20 answers)*

- Presentation of projects, open discussion about specific technologies.
- participation in events
- nothing is available at the moment
- Organization of trainings, workshops and panels
- Sharing our experiences, software and data + data models
- Participation in events, sharing resources/knowledge
- Hosting events and liaison with other groups
- Share practice and bad/good experiences working on e-culture project and implementing usage of AI
- Our experience
- We organize Annif-tutorials regularly, we could also organize a webinar about our efforts in AI and automated cataloguing.
- Workshops, Webinars, Examples
- Sharing in-house R&D experiences
- Organization of events, region-related research and survey distribution...
- Our current work on AV-collections
- Presentations, maybe hosting an event
- Having large funds of manuscripts and old prints (mainly Cyrillic), our library could contribute to the improvement of text and image recognition mechanisms by AI

- Knowledge sharing; participation in events and networking
- Small scale research projects, technology watch, later also data models
- I could contribute by sharing some insights from our AI experiments in cataloguing. Additionally, I could participate in some discussions or collaborative events.