ELSA: European Lighthouse on Secure and Safe AI

Coordinator:
Prof. Dr. Mario Fritz
CISPA Helmholtz Center for Information Security

Presenter: Prof. Plamen Angelov, WP3 Lead
Lancaster University, UK

Duration:
September 2022 to August 2026

Web: https://elsa-ai.eu
Twitter: @elsa_lighthouse

Founding members:

CISPA HELMHOLTZ CENTER FOR INFORMATION SECURITY
UNIVERSITY OF HELSINKI
KTH ROYAL INSTITUTE OF TECHNOLOGY
COMPUTER VISION CENTER
ITALIAN INSTITUTE OF TECHNOLOGY
VALEO.AI
UNIVERSITY OF BIRMINGHAM
THE ALAN TURING INSTITUTE
UNIVERSITY OF MODENA AND REGGIO EMILIA
NATIONAL INTERUNIVERSITY CONSORTIUM FOR INFORMATICS
UNIVERSITY OF CAGLIARI
UNIVERSITY OF GENOA
UNIVERSITY OF MILANO
POLYTECHNIC UNIVERSITY OF TURIN
PAL ROBOTICS
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
ETH ZURICH
NVIDIA
UNIVERSITY OF OXFORD
LANCASTER UNIVERSITY
MAX PLANCK SOCIETY
NATIONAL INSTITUTE FOR RESEARCH IN COMPUTER SCIENCE AND AUTOMATION
EUROPEAN MOLECULAR BIOLOGY LABORATORY
PLURIBUS ONE
YOOZ
LEONARDO

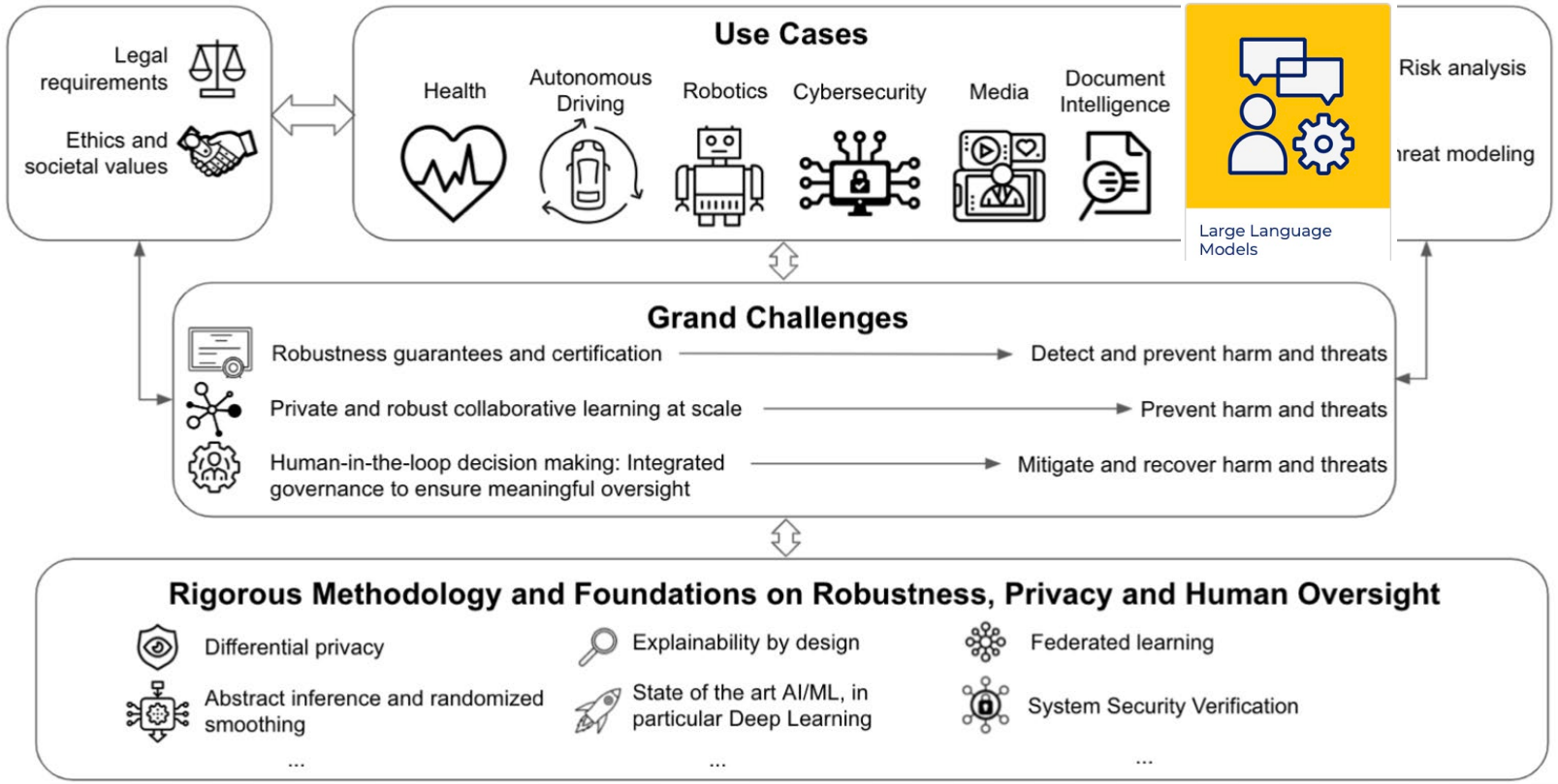
ellis

elsa

Funded by the European Union

European Network of Excellence: European Lighthouse on Secure and Safe AI (ELSA)

We are creating a **community** to build methods that address **needs of industry**, based on **solid theoretical foundations** to comply and innovate in our **European** understanding of **Trustworthy, Secure, and Safe AI.**



https://elsa-ai.eu

CISPA HELMHOLTZ CENTER FOR INFORMATION SECURITY
UNIVERSITY OF HELSINKI
KTH ROYAL INSTITUTE OF TECHNOLOGY
COMPUTER VISION CENTER
ITALIAN INSTITUTE OF TECHNOLOGY
VALEO.AI
UNIVERSITY OF MODENA AND REGGIO EMILIA
NATIONAL INTERUNIVERSITY CONSORTIUM FOR INFORMATICS
UNIVERSITY OF CAGLIARI
UNIVERSITY OF GENOA
UNIVERSITY OF MILANO
POLYTECHNIC UNIVERSITY OF TURIN
MAX PLANCK SOCIETY
NATIONAL INSTITUTE FOR RESEARCH IN COMPUTER SCIENCE AND AUTOMATION
EUROPEAN MOLECULAR BIOLOGY LABORATORY
PLURIBUS ONE
YOOZ
LEONARDO
PAL ROBOTICS
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
ETH ZURICH
NVIDIA
UNIVERSITY OF OXFORD
LANCASTER UNIVERSITY
UNIVERSITY OF BIRMINGHAM
THE ALAN TURING INSTITUTE

Grand Challenges

- Certifiable robustness
- Private, federated learning
- Human in the loop decision making

Use Case Owners

Academia *Industry*

Use Cases

- **Health**
  Federated Genome Medicine
- **Autonomous Driving**
  Robust Perception
- **Robotics**
  Learning Through Human Interaction
- **Media Analytics**
  Tackling Disinformation
- **Cybersecurity**
  Malware Detection
- **Document Intelligence**
  Document VQA
- **LLM**
  LLM capture the flag

ELSA Impact

Science

- >200 open-access scientific publications, 80% top tier venues
- >300 applications in ELSA topics to PhD program
- Software libraries/tool
- ELSA Benchmarks Platform

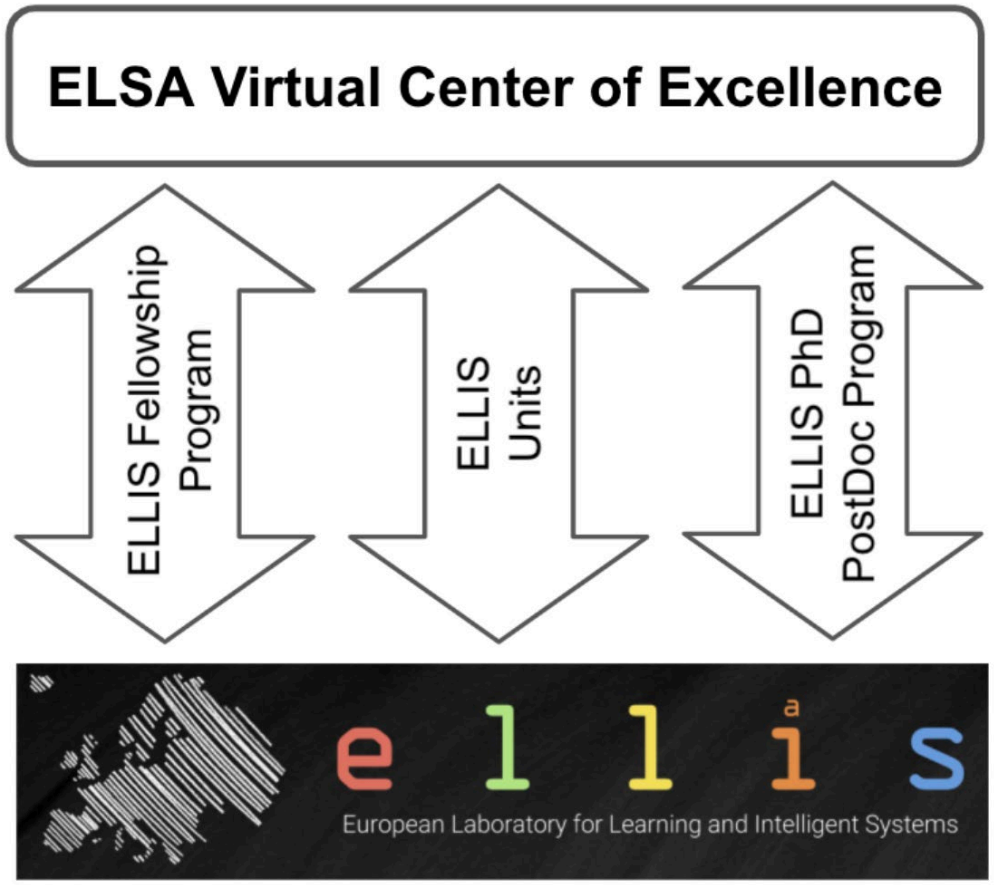
Society

- Open Letter, House of Lords
- AI Ethics Roundtable on Human-Machine Interaction
- Public Outreach
- Press & Social Media

Economy

- Use Case co-ownership
- ELSA Industry Open Call
- Mentoring start-ups

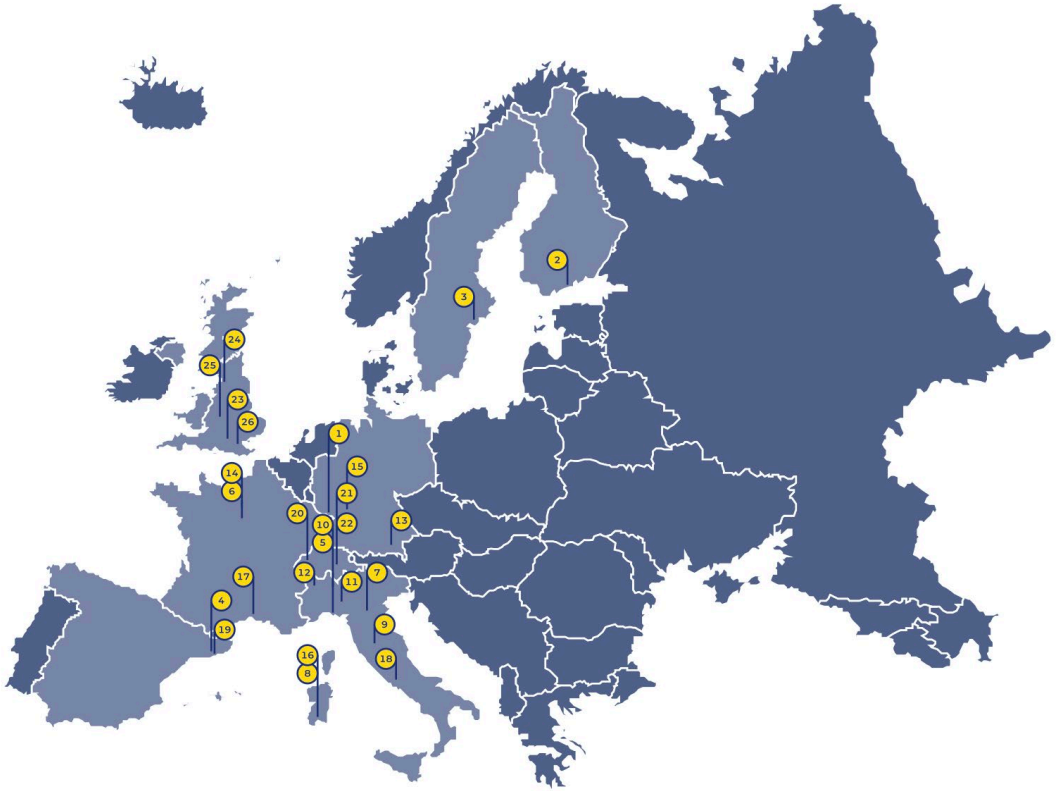
European Network of Excellence

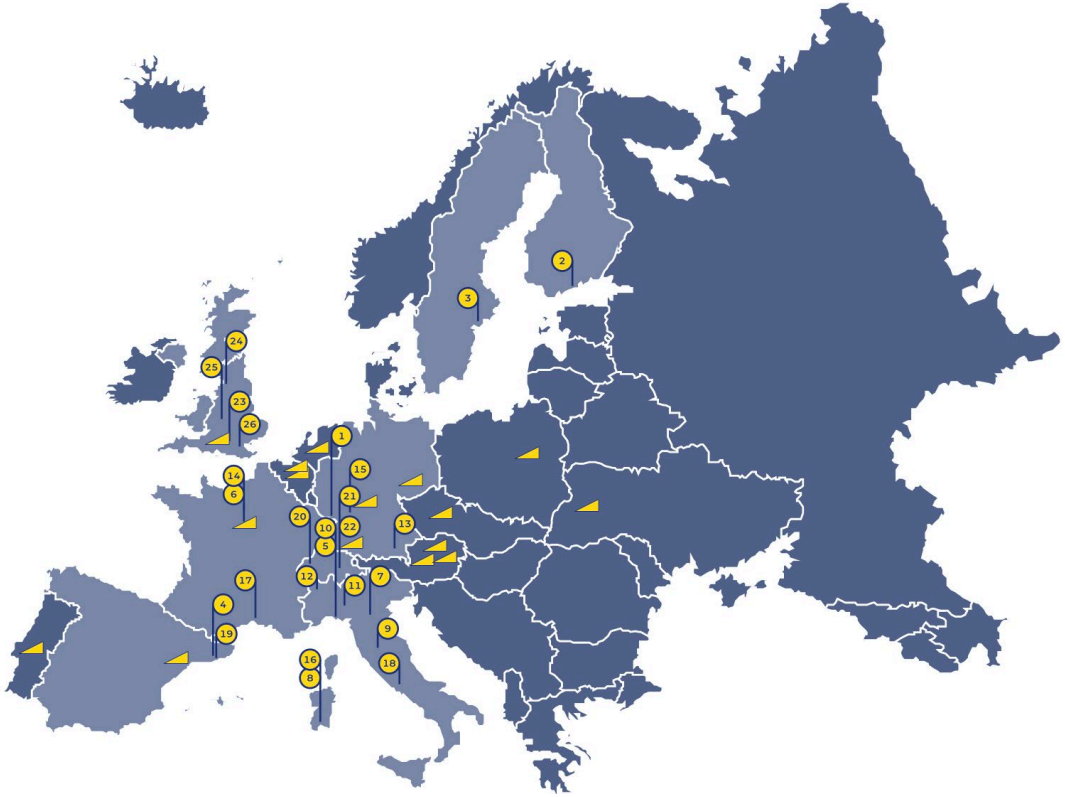


Common Understanding of Excellence

Internal Review
Units, Programs
Scholars, Fellows

European Network
Founding Members to ELLIS Units





EU – ELLIS - ELSA

elise
European Network of AI Excellence Centres

elsa
European Lighthouse on Secure and Safe AI

ZUSE SCHOOL ELIZA

ELIAS

ELLIOT

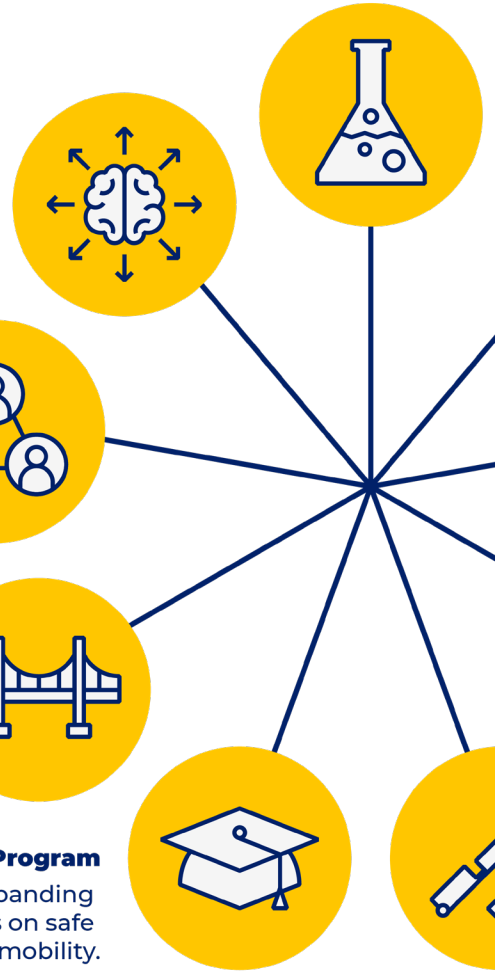
ELLE
European Doctoral Network on Secure and Safe Artificial Intelligence in the Age of Foundation Models

ellis

KEY ACTIVITIES

Research Program

ELSA will form Research Programmes around foundational methodology for safe and secure machine-learning-based AI that provide a nucleus for key innovations.



Spread knowledge

ELSA will spread knowledge and a European vision of AI across stakeholders and general public.

Strategic Research Agenda

ELSA will build and drive forward a Strategic Research Agenda in safe and secure AI.

Network

ELSA is designed to grow a network of excellence for safe and secure AI based on common standards.

Benchmarks

ELSA will define benchmarks to steer innovation towards addressing grand challenges that hamper deployment.

Bridge Head Office

ELSA's bridge head office will leverage momentum and synergies as well as avoid duplication across networks.

Innovation Lab

The ELSA innovation lab will provide a platform for driving and exploiting socially-beneficial innovation.

PhD/PostDoc Program

ELSA will acquire talent by expanding the ELLIS PhD/Postdoc program to topics on safe and secure AI as well as support for mobility.

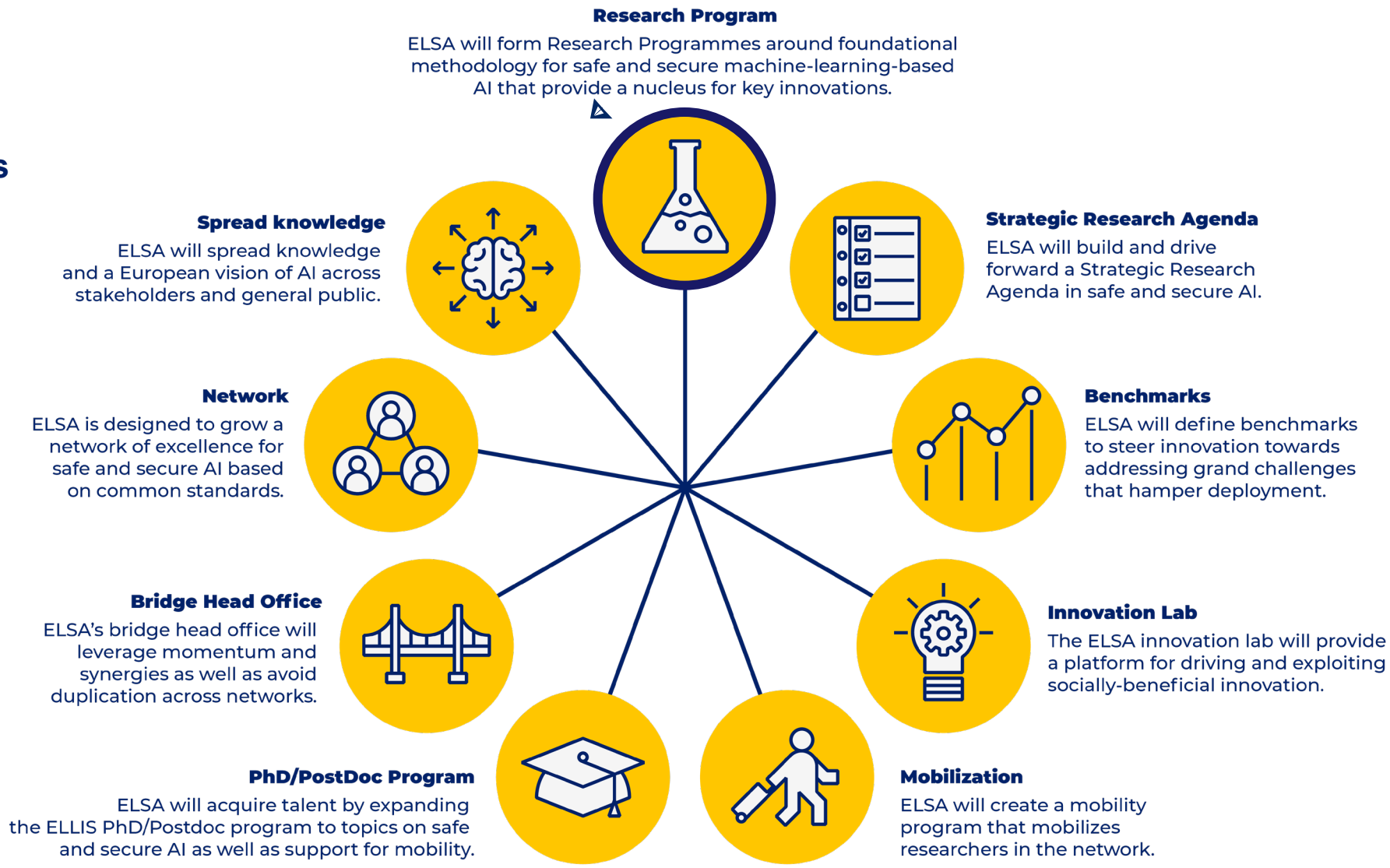
Mobilization

ELSA will create a mobility program that mobilizes researchers in the network.

KEY ACTIVITIES

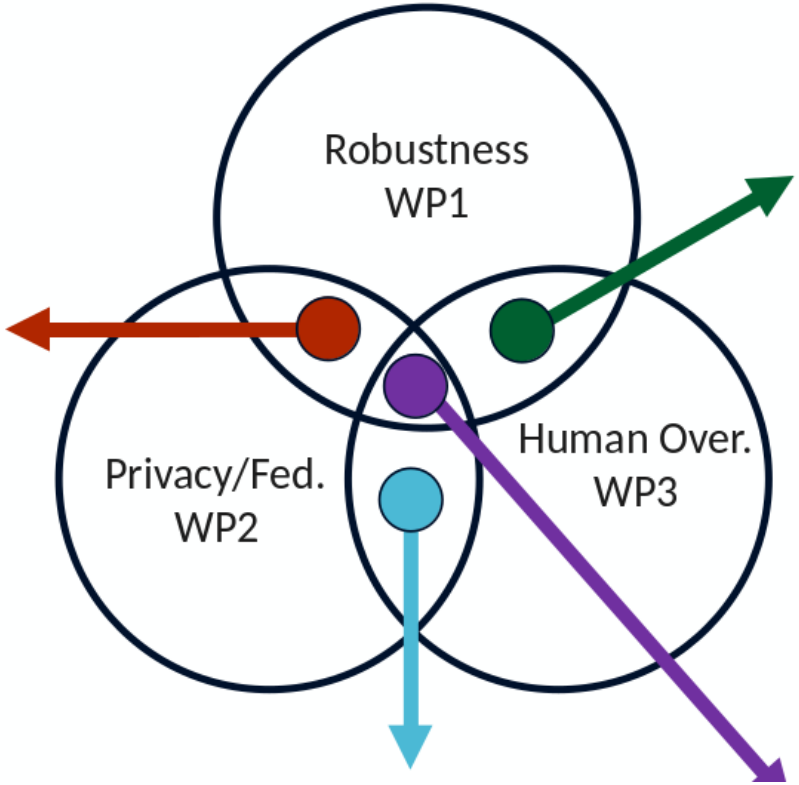
- **>200 open-access scientific publications, 80% top tier venues**
- **Top researcher in ML (virtual Csranking)**
- **Cross Sites**
- **Cross Topic**

#	Institution	Count	Faculty
1	ELSA	158.6	24
2	Technion	143.1	43
3	EPFL	136.7	34
4	ETH Zurich	124.7	32
5	Tel Aviv University	92.5	21
6	University of Cambridge	91.1	21
7	University of Tübingen	76.9	24
8	Max Planck Society	74.4	17
9	University of Edinburgh	71.0	47
10	TU Munich	64.9	29



Examples of Cross-Topic Contributions

- Dataset and Lessons Learned from the 2024 SaTML LLM Capture-the-Flag Competition
  Debenedetti; Rando; Paleka; Florin; Albastroiu; Cohen; Lemberg; Ghosh; Wen; Salem; Cherubin; Zanella-Beguelin; Schmid; Klemm; Miki; Li; Kraft; Fritz; Tramèr; Abdelnabi; Schönherr
  In: Neural Information Processing Systems (NeurIPS), 2024.
- + LLM ELSA Benchmark/Challenge



- On the Robustness of Adversarial Training Against Uncertainty Attacks
  Ledda, Emanuele; Scodeller, Giovanni; Angioni, Daniele; Piras, Giorgio; Cinà, Antonio Emanuele; Fumera, Giorgio; Biggio, Battista; Roli, Fabio
  Journal Article
  In: Pattern Recognition, vol. 172, pp. 112519, 2026
- Pixel-level Certified Explanations via Randomized Smoothing
  Anani; Lorenz; Fritz; Schiele
  In: International Conference on Machine Learning (ICML), 2025.

- On consistent Bayesian inference from synthetic data
  Räisä, Jälkö, Honkela; Journal of Machine Learning Research (JMLR), 2024
- Noise-aware differentially private regression via meta-learning
  Räisä, Markou, Ashman, Bruinsma, Tobaben, Honkela, Turner; In: Neural Information Processing Systems (NeurIPS), 2024

- Causality Is Key to Understand and Balance Multiple Goals in Trustworthy ML and Foundation Models
  Binkyte; Sheth; Jin; Havaei; Schölkopf; Fritz
  ArXiv'25
- ELSA SRA

KEY ACTIVITIES

- **ELSA Strategic Research Agenda for Secure and Safe AI in Europe**
- **Contribution to SRIDA of ADRA**
- **Both to be updated 2025**

Research Program

ELSA will form Research Programmes around foundational methodology for safe and secure machine-learning-based AI that provide a nucleus for key innovations.

Spread knowledge

ELSA will spread knowledge and a European vision of AI across stakeholders and general public.

Strategic Research Agenda

ELSA will build and drive forward a Strategic Research Agenda in safe and secure AI.

Network

ELSA is designed to grow a network of excellence for safe and secure AI based on common standards.

Benchmarks

ELSA will define benchmarks to steer innovation towards addressing grand challenges that hamper deployment.

Bridge Head Office

ELSA's bridge head office will leverage momentum and synergies as well as avoid duplication across networks.

Innovation Lab

The ELSA innovation lab will provide a platform for driving and exploiting socially-beneficial innovation.

PhD/PostDoc Program

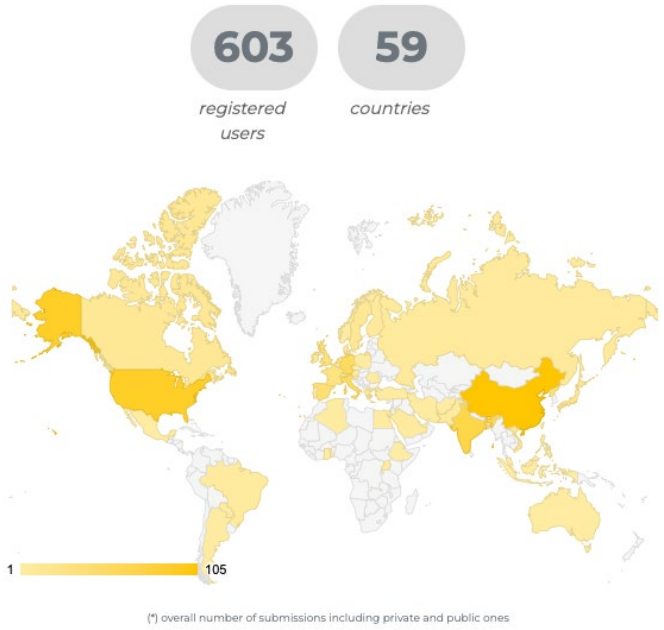
ELSA will acquire talent by expanding the ELLIS PhD/Postdoc program to topics on safe and secure AI as well as support for mobility.

Mobilization

ELSA will create a mobility program that mobilizes researchers in the network.

KEY ACTIVITIES

- **Internationally successful benchmarking platform for Secure and Safe AI**



Research Program

ELSA will form Research Programmes around foundational methodology for safe and secure machine-learning-based AI that provide a nucleus for key innovations.

Spread knowledge

ELSA will spread knowledge and a European vision of AI across stakeholders and general public.

Strategic Research Agenda

ELSA will build and drive forward a Strategic Research Agenda in safe and secure AI.

Network

ELSA is designed to grow a network of excellence for safe and secure AI based on common standards.

Benchmarks

ELSA will define benchmarks to steer innovation towards addressing grand challenges that hamper deployment.

Bridge Head Office

ELSA's bridge head office will leverage momentum and synergies as well as avoid duplication across networks.

Innovation Lab

The ELSA innovation lab will provide a platform for driving and exploiting socially-beneficial innovation.

PhD/PostDoc Program

ELSA will acquire talent by expanding the ELLIS PhD/Postdoc program to topics on safe and secure AI as well as support for mobility.

Mobilization

ELSA will create a mobility program that mobilizes researchers in the network.

ELSA BENCHMARKS PLATFORM

URL: https://benchmarks.elsa-ai.eu/

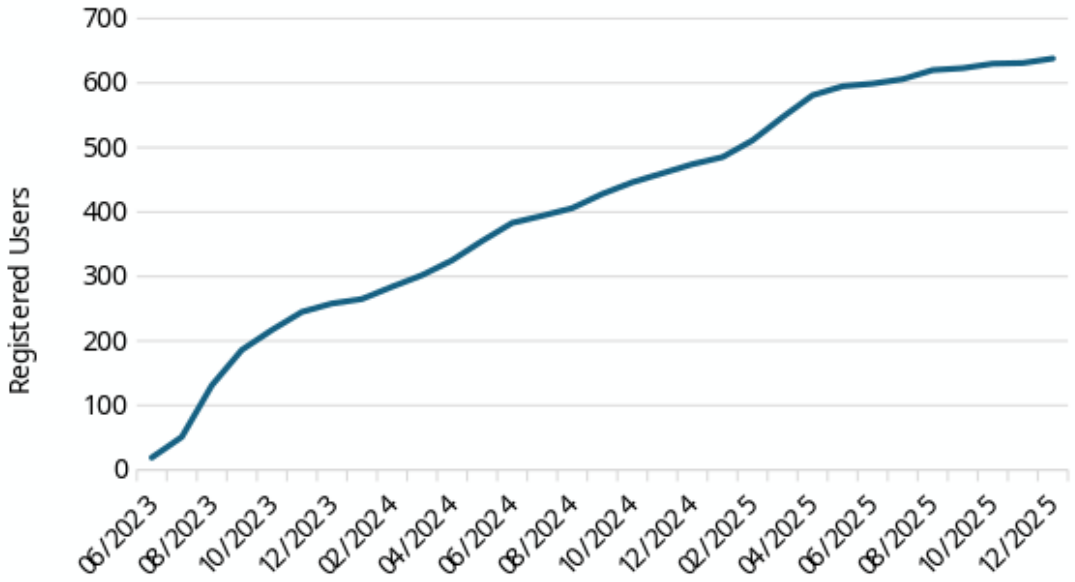
- Public ELSA resources
- Competition hosting
- Automatic evaluation of submissions and leaderboards
- New functionality implemented in this period to support better Blue / Red team schemes (e.g. Health challenges)

FACTS AND FIGURES

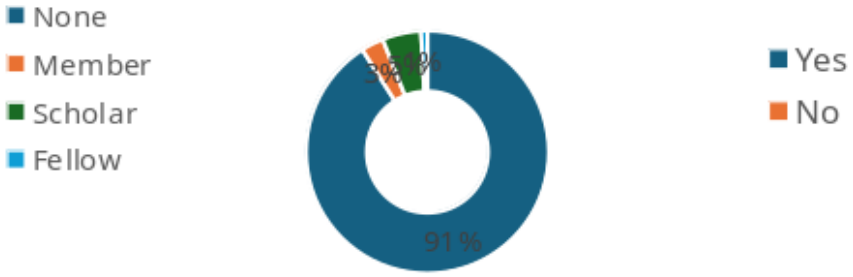
638
Registered Users

4152 Visits

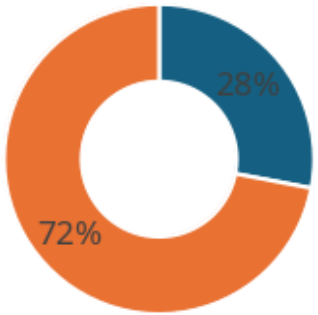
59 Countries

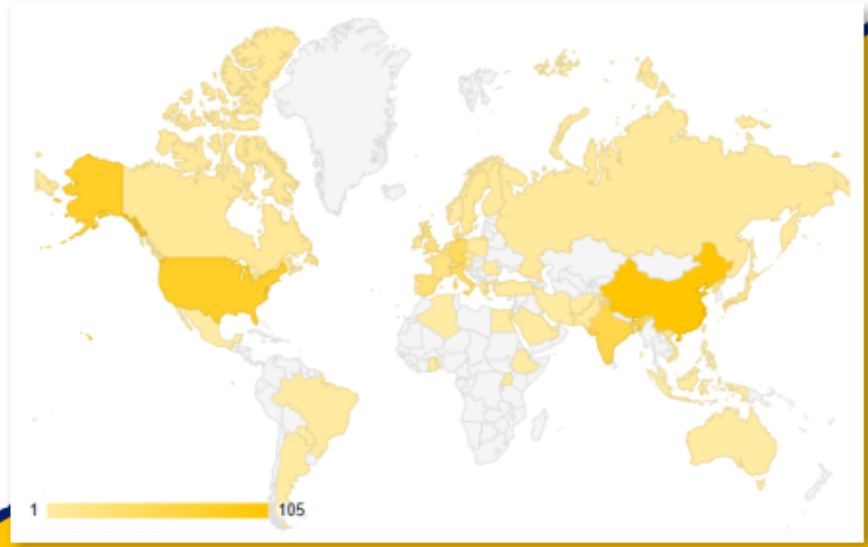


ELLIS Members (new registrations)



ELSA Members





FACTS AND FIGURES

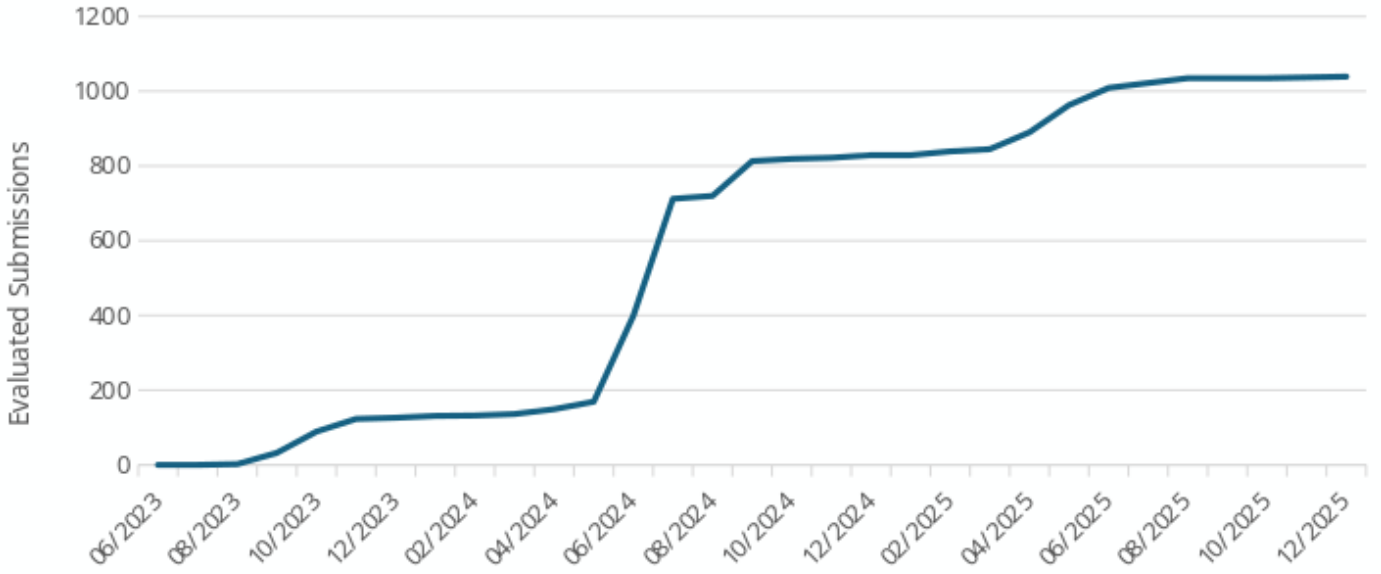
20 Competitions organised

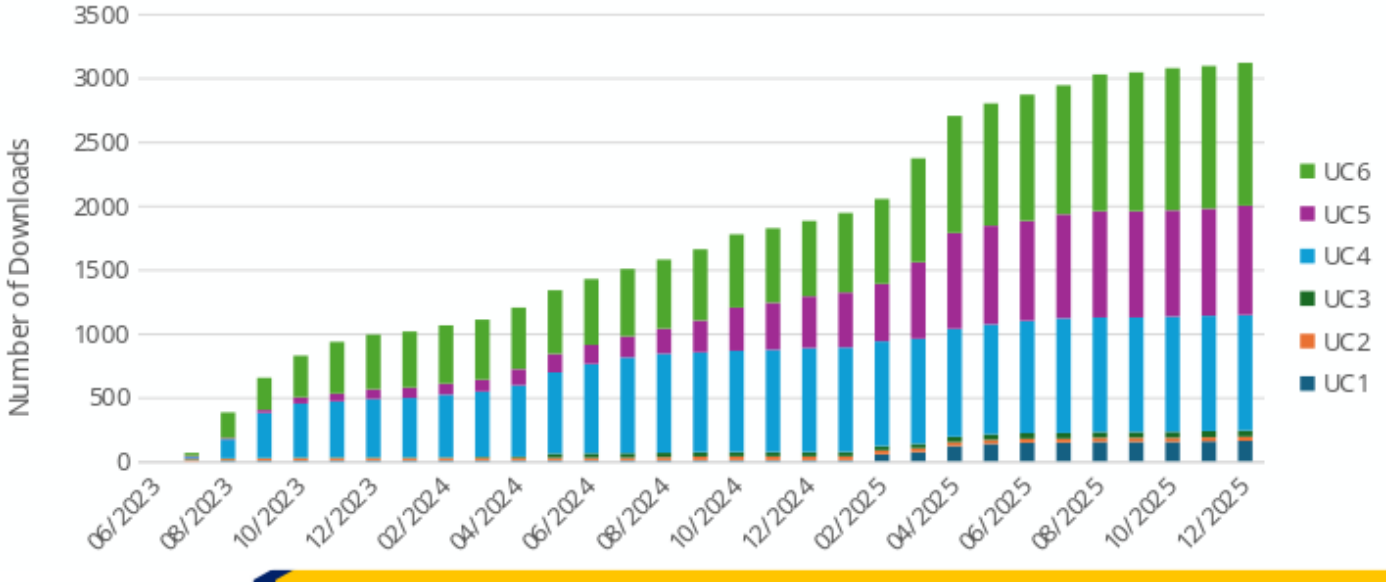
21 Events organised in key venues

1041 Evaluated Methods

2968 Downloads

Note: UC2 is underrepresented as competitions were hosted outside the ELSA benchmarks platform





KEY ACTIVITIES

- **2 successful Industry Call supporting and mentoring startups and SMEs**




Starlab LIVING SCIENCE



Research Program

ELSA will form Research Programmes around foundational methodology for safe and secure machine-learning-based AI that provide a nucleus for key innovations.

Spread knowledge

ELSA will spread knowledge and a European vision of AI across stakeholders and general public.

Strategic Research Agenda

ELSA will build and drive forward a Strategic Research Agenda in safe and secure AI.

Network

ELSA is designed to grow a network of excellence for safe and secure AI based on common standards.

Benchmarks

ELSA will define benchmarks to steer innovation towards addressing grand challenges that hamper deployment.

Bridge Head Office

's bridge head office will everage momentum and synergies as well as avoid duplication across networks.

Innovation Lab

The ELSA innovation lab will provide a platform for driving and exploiting socially-beneficial innovation.

PhD/PostDoc Program

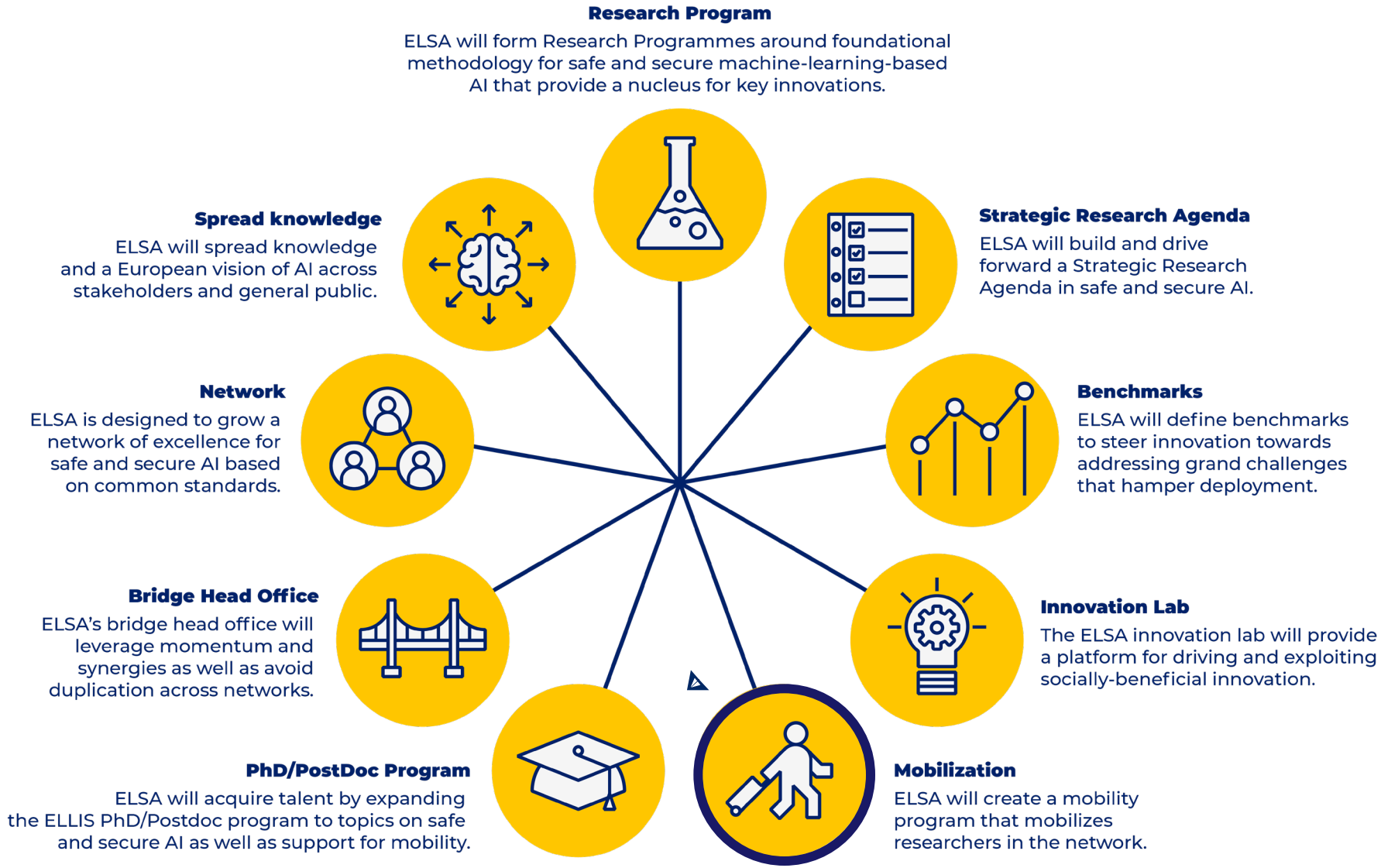
ELSA will acquire talent by expanding the ELLIS PhD/Postdoc program to topics on safe and secure AI as well as support for mobility.

Mobilization

ELSA will create a mobility program that mobilizes researchers in the network.

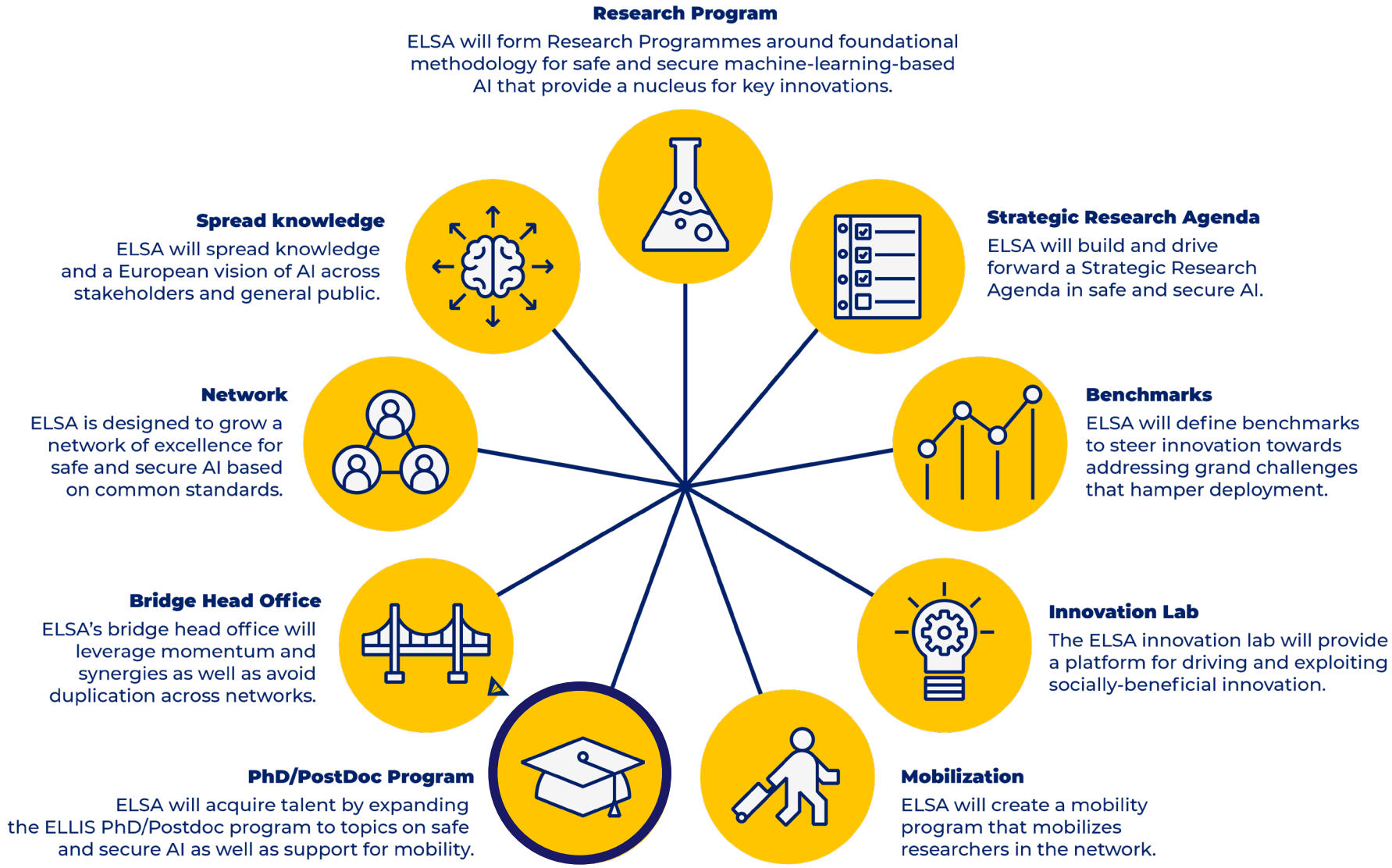
KEY ACTIVITIES

- Scientific Exchange
- Summerschools
- Doctoral Symposia
- Lectures
- Workshops
- Assemblies



KEY ACTIVITIES

- **Close collaboration with ELLIS PhD and PostDoc program**
  - 3200 registrations
  - 1500 applicants
  - Over 1/3 in ELSA topics
  - 2/3 Europe



KEY ACTIVITIES

- **Workshop and events**
- **Peer-reviewed workshops co-located at top international venues**

Research Program

ELSA will form Research Programmes around foundational methodology for safe and secure machine-learning-based AI that provide a nucleus for key innovations.

Spread knowledge

ELSA will spread knowledge and a European vision of AI across stakeholders and general public.

Strategic Research Agenda

ELSA will build and drive forward a Strategic Research Agenda in safe and secure AI.

Network

ELSA is designed to grow a network of excellence for safe and secure AI based on common standards.

Benchmarks

ELSA will define benchmarks to steer innovation towards addressing grand challenges that hamper deployment.

Bridge Head Office

ELSA's bridge head office will leverage momentum and synergies as well as avoid duplication across networks.

Innovation Lab

The ELSA innovation lab will provide a platform for driving and exploiting socially-beneficial innovation.

PhD/PostDoc Program

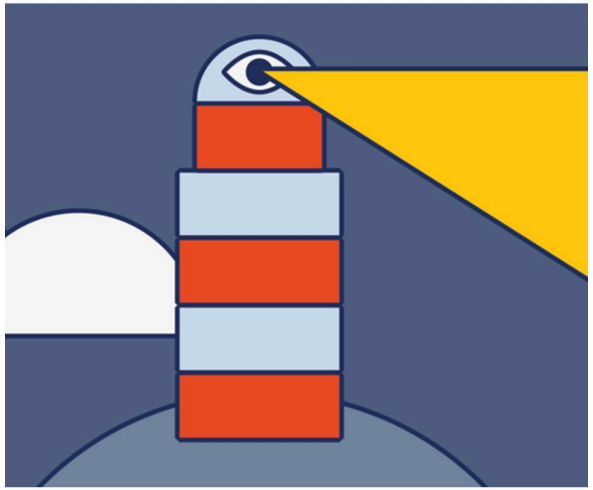
ELSA will acquire talent by expanding the ELLIS PhD/Postdoc program to topics on safe and secure AI as well as support for mobility.

Mobilization

ELSA will create a mobility program that mobilizes researchers in the network.

KEY ACTIVITIES

- Lectures, Tutorials, Courses
- Summerschool
- Outreach
- Web, LinkedIn, Twitter
- >60 Software and Tools



Research Program

ELSA will form Research Programmes around foundational methodology for safe and secure machine-learning-based AI that provide a nucleus for key innovations.

Spread knowledge

ELSA will spread knowledge and a European vision of AI across stakeholders and general public.

Strategic Research Agenda

ELSA will build and drive forward a Strategic Research Agenda in safe and secure AI.

Network

ELSA is designed to grow a network of excellence for safe and secure AI based on common standards.

Benchmarks

ELSA will define benchmarks to steer innovation towards addressing grand challenges that hamper deployment.

Bridge Head Office

ELSA's bridge head office will leverage momentum and synergies as well as avoid duplication across networks.

Innovation Lab

The ELSA innovation lab will provide a platform for driving and exploiting socially-beneficial innovation.

PhD/PostDoc Program

ELSA will acquire talent by expanding the ELLIS PhD/Postdoc program to topics on safe and secure AI as well as support for mobility.

Mobilization

ELSA will create a mobility program that mobilizes researchers in the network.

Policy Impact

- Independent expert study on technical means for implementing article 50 of EU AI Act towards the code of practice
- Panel on "Loss of Control" at EU AI Office Workshop on "Towards defining best practices for systemic risk evaluations"
- Talk on "Fundamental Risks in the Current Deployment of General-Purpose AI Models: What Have We (Not) Learnt From Cybersecurity?" at European AI Office workshop “Evaluating General-Purpose AI Models with Systemic Risk”
- Expert opinion for evidence review report “Successful and timely uptake of artificial intelligence in science in the EU”
- "AI, Data, and Robotics ‘made in Europe’: Research agendas from the EU AI and robotics Networks of Excellence”
- Uplift study together with SaferAI on impact of AI on Cybersecurity Risks

Society / Boostrapping Effect

- Advisory Board for Swedish Foundation for Strategic Research (SSF)

  Multidisciplinary Research Centre in Cyber-Resilience for AI-systems (MRC CRAI)
- Advisory Board for French

  Cluster SequoIA: Security and Confidence in and by AI digital systems
- Advisory Board

  European frontier AI initiative

- Projects:
  - HORIZON-CL4-2024-HUMAN-03 ELLIOT: Multimodal Generalist Foundation Models (MGFMs)
  - HORIZON-CL4-2025-04-DIGITAL-EMERGING-07: under review

- Submission to Marie Curie Program (MSCA): ELLE - negotiations
- German Ministry for Research and Education: in preparation

Outlook

- Summerschool on “Responsible and Trustworthy AI in Drug Discovery” at CISPA/Saarland
- European AI and Cybersecurity Hackathon Championships
- Information and awarenss of challenges in AI & Cybersecurity
  - Digital Innovation Hub DAISEC
  - Workshop format
- New PhD Innovation track focus on startups
- Forecasting activities
- European Frontier AI Initiative
- Final general assembly

ELSA: European Lighthouse on Secure and Safe AI

Coordinator:
Prof. Dr. Mario Fritz
CISPA Helmholtz Center for Information Security

Duration:
September 2022 to August 2026

Web: https://elsa-ai.eu
Twitter: @elsa_lighthouse

Founding members:

CISPA HELMHOLTZ CENTER FOR INFORMATION SECURITY
UNIVERSITY OF HELSINKI
KTH ROYAL INSTITUTE OF TECHNOLOGY
COMPUTER VISION CENTER
ITALIAN INSTITUTE OF TECHNOLOGY
VALEO.AI
UNIVERSITY OF BIRMINGHAM
THE ALAN TURING INSTITUTE
UNIVERSITY OF MODENA AND REGGIO EMILIA
NATIONAL INTERUNIVERSITY CONSORTIUM FOR INFORMATICS
UNIVERSITY OF CAGLIARI
UNIVERSITY OF GENOA
UNIVERSITY OF MILANO
POLYTECHNIC UNIVERSITY OF TURIN
PAL ROBOTICS
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
ETH ZURICH
NVIDIA
UNIVERSITY OF OXFORD
LANCASTER UNIVERSITY
MAX PLANCK SOCIETY
NATIONAL INSTITUTE FOR RESEARCH IN COMPUTER SCIENCE AND AUTOMATION
EUROPEAN MOLECULAR BIOLOGY LABORATORY
PLURIBUS ONE
YOOZ
LEONARDO

ellis

elsa

Funded by the European Union