Artificial intelligence, Data and Robotics ecosystem

https://adra-e.eu/

Call: A human-centred and ethical development of digital and industrial technologies 2021
Topic: Horizon-CL4-2021-Human-01
Type of action: Coordination and Support actions
Grant agreement Nº: 101070336

WP Nº3:	**ADR Awareness Centre**
Deliverable Nº3.4:	**Report on ADR Awareness Centre activities and awareness day Period 2**
Lead partner:	**University of Galway**
Version Nº:	**1.0**
Date:	**30/06/2025**
Dissemination level[1]:	**PU**

[1] **PU:** Public**; CO:** Confidential, only for members of the consortium (including the Commission Services)

Document information	
Deliverable Nº and title:	D3.4 – Period 2 report on ADR Awareness Centre activities
Version Nº:	1.0
Lead beneficiary:	University of Galway (UoG)
Author(s):	Fatemeh Ahmadi Zeleti (UoG), Heike Vornhagen (UoG)
Reviewers:	Sonja Zillner (Siemens)
Submission date:	30/06/2025
Due date:	30/06/2025
Type[2]:	R
Dissemination level[3]:	PU

Document history			
Date	Version	Author(s)	Comments
01/04/2025	0.1	Fatemeh Ahmadi Zeleti	Table of Content, Format, structure, draft 1
30/05/2025	0.2	Heike Vornhagen	Drafted Book 1 and 2025 Awareness Day subsections
03/06/2024	0.3	Fatemeh Ahmadi Zeleti	Revised and reformatted sections 1, 2 and 3
04/06/2025	0.4	Fatemeh Ahmadi Zeleti	Restructure and revised section 4 and 5 Revised "outcome" & "Statistics on Attendees" for 2025 event
05/06/2025	0.5	Fatemeh Ahmadi Zeleti	SRIDA alignment analysis completed and inserted
06/06/2025	0.6	Fatemeh Ahmadi Zeleti	Final check and submission for internal review
16/06/2025	0.7	Sonja Zillner	Review
23/06/2025	1.0	Fatemeh Ahmadi Zeleti	Final version

[2] **R**: Report, **DEC:** Websites, patent filling, videos; **DEM:** Demonstrator, pilot, prototype; **OTHER:** Software Tools
[3] **PU:** Public**; CO:** Confidential, only for members of the consortium (including the Commission Services)

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of the Adra-e consortium and can in no way be taken to reflect the views of the European Union.

The European Union is established in accordance with the Treaty on European Union (Maastricht). There are currently 28 Member States of the Union. It is based on the European Communities and the Member States cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice and the Court of Auditors (http://europa.eu/).

Adra-e has received funding from the European Union's Horizon Europe under grant agreement 101070336.

Table of Contents

Executive Summary

This deliverable outlines the progress of Work Package 3 (WP3) of the Adra-e project during Period 2 (M18–M36), focusing on two core tasks: the ADR Awareness Centre (T3.1) and the ADR Awareness Days (T3.4).

The ADR Awareness Centre has become a central hub for accessible, high-quality educational resources on AI, Data, and Robotics (ADR), supporting diverse stakeholders including researchers, SMEs, and policymakers. With over 100 curated materials, API integration with the AI-on-Demand platform, and a transition to the ADRA Association, the Centre ensures long-term sustainability and EU-wide impact.

A key advancement during this period was the development of a resource taxonomy strategically aligned with the SRIDA framework, allowing for insight into how community contributions reflect EU priorities. Analysis revealed strong representation in technical and data-focused areas, with moderate attention to policy and trust, and fewer resources in areas such as skills, deployment, and fundamental rights—highlighting important opportunities for future focus.

The two ADR Awareness Day events held in 2024 and 2025 attracted over 100 participants from 16 countries. Delivered virtually, they featured expert panels on trustworthy AI, generative technologies, and human–robot interaction, serving as effective dissemination platforms for WP3 outputs.

Together, these activities have strengthened ADR knowledge-sharing, supported EU policy alignment, and provided strategic visibility into where further community engagement is needed. WP3 continues to deliver long-term value by promoting transparency, collaboration, and responsible innovation across the European ADR ecosystem.

1. Introduction

The European AI, Data, and Robotics (ADR) ecosystem is vast and rapidly evolving. Its holistic nature and the diverse historical backgrounds of its stakeholders make it inherently complex. This complexity is further amplified by the sheer scale of activities and the multitude of resources and materials that must be considered. Consequently, a one-size-fits-all approach is impractical, and comprehensive engagement and collaboration across the ecosystem are essential, particularly in education and the provision of resources.

To address these challenges, Tasks T3.1 and T3.3 of Work Package 3 (WP3) have been developed. These tasks aim to support stakeholders by providing educational materials and raising awareness about the trustworthiness of AI solutions for consumers.

- **T3.1 ADR Awareness Centre**

  The primary objective of this task is to establish a structured framework for outreach and awareness of ADR, fostering collaboration and alignment among all projects within the Partnership, as well as with relevant external initiatives. This task also includes the development of two Open Access books that capture critical outputs from the projects.

- **T3.4 ADR Awareness Day**

  The aim of this task is to promote WP3 and disseminate the materials developed within it through a series of ADR Awareness Day seminars. These seminars will showcase the ADR Awareness Centre, highlight findings on AI externalities (both good and bad practices), and present the selected AI Trust Label along with its selection process and methodology. The ADR Awareness Day will target society at large and will address the ethical and trustworthy use of AI. Key outreach groups include:

  a) Public/Citizens
  b) Businesses (Large/SMEs)
  c) Public Administrators (Large/SMEs)
  d) Educators

  The goal is to increase awareness of the acceptability and trustworthiness of ADR across these groups.

1.1. Target group of the Deliverable

The objective of this deliverable (D3.4) of the Adra-e project is to provide a comprehensive report on the activities carried out and the progress achieved in Tasks T3.1 (ADR Awareness Centre) and T3.4 (ADR Awareness Day) during Period 2 (M18–M36). This report highlights key milestones, outputs, and the overall impact of these tasks on the broader project objectives.

The target audience for this deliverable includes:

- The five Adra funding organizations, who require insights into the outcomes and impacts of the project.
- The Adra-e Project Management team, for monitoring progress and aligning with project objectives.
- Adra-e Consortium members, who will benefit from understanding how these tasks contribute to the collective efforts and collaboration within the consortium.
- Adra-e Advisory Board members, who provide expert guidance and strategic direction based on the progress reported.

- Adra members, who will gain insights into the educational and outreach activities aimed at strengthening the ADR ecosystem.

By providing a clear and detailed account of these activities and their outcomes, this deliverable ensures transparency, facilitates informed decision-making, and strengthens the collaborative efforts within the Adra-e project.

1.2. Document Outline

This deliverable is structured into six main sections, each detailing key aspects of Work Package 3 (WP3) activities during the second reporting period (M18–M36).

- Section 1 – Introduction: Outlines the context, purpose, and target audience of the deliverable.
- Section 2 – Collaboration and Impact: Summarizes key collaborations established across the Adra-e ecosystem and the expected impacts resulting from the completion of Tasks 3.1 and 3.4.
- Section 3 – ADR Awareness Centre: Provides a detailed overview of the Awareness Centre's vision, resource strengths, usage metrics, sustainability plans, integration with AI-on-Demand, and performance monitoring framework.
- Section 4 – ADR Awareness Days: Presents the objectives, delivery, and impact of the two major stakeholder engagement events held in 2024 and 2025.
- Section 5 – Dissemination of ADR Content through Open Access Books: Describes the editorial process and thematic scope of two books developed as part of WP3 dissemination efforts, including summaries of submitted abstracts.
- Section 6 – Conclusion: Concludes the deliverable by reflecting on WP3 achievements and the long-term value of the activities delivered.

An appendix provides the full list of abstracts submitted for Book 1 and Book 2.

2. Collaboration and Impact

This section outlines the key collaborations and impacts of T3.1 and T3.4 in period 2.

2.1. Key Collaborations

Transforming the ADR Skills and Education landscape in Europe requires strong collaboration with European stakeholders, initiatives, and both past and ongoing EC-funded projects. A central principle of WP3 is to build on these relationships to leverage existing knowledge, skills, and materials within the ADR ecosystem. During Period 1, WP3 focused on several key collaborations:

- **Collaboration with WP1**
  WP3 worked closely with WP1 to enhance the visibility of the ADR Awareness Centre. This involved crafting a standard message to be disseminated across partnership projects. WP1's objective to foster a dynamic ADR partnership through synergy creation and sharing of best practices directly supports WP3's outreach goals.

- **Collaboration with AI-for-Europe**
  The AI-for-Europe project is particularly linked to WP3 and Task 3.1 (ADR Awareness Centre). As a CSA project funded by the European Commission, AI-for-Europe is centered on developing the community-driven AI-on-Demand platform to empower European research and innovation in Artificial Intelligence, while ensuring the European values of quality, trustworthiness, and explainability. Multiple meetings—both virtual and in-person—have been held between Adra-e and AI-for-Europe to explore data-sharing opportunities between the AI-on-Demand platform and the ADR Awareness Centre. These discussions remain ongoing. WP3 is committed to supporting any data-sharing decisions and activities that arise, further strengthening this collaboration to promote the ADR Awareness Centre and its resources.

- **Collaborations with Experts for Edited Books**
  In the context of Task 3.1, WP3 has collaborated with several experts from across Europe in the development of two edited volumes during this period. These collaborations have contributed significant expertise and insights to the materials produced.

- **Collaboration via the ADR Awareness Days**
  As part of Task 3.4, WP3 organized two ADR Awareness Day events—one in 2024 and another in 2025—designed as public and virtual gatherings. These events served as vital platforms to promote trust and excellence in the ADR ecosystem and showcased the collaborative and inclusive spirit that underpins WP3's activities.

  Each ADR Awareness Day featured a diverse range of expert speakers and panel members drawn from across Europe. These experts, representing leading research institutions, industry, public sector organizations, and civil society, shared their insights on emerging trends, challenges, and opportunities in AI, Data, and Robotics. Discussions spanned topics such as ethical AI, trustworthy and explainable systems, regulatory frameworks, skills and education in ADR, and collaborative innovation in the European context.

In addition to showcasing WP3 outcomes—including the ADR Awareness Centre and resources—the events provided an opportunity for deep, cross-sectoral dialogue and co-creation. Panel discussions and interactive sessions fostered exchange of ideas and best practices, encouraging participants to learn from one another and explore new ways to build trust in ADR systems.

Notable highlights include:

- **2024 ADR Awareness Day** – Held virtually, this event convened a broad audience from across Europe to build trust and excellence through partnership in ADR. The event details and recordings are available here: https://adra-e.eu/events/adr-partnership-awareness-day-building-trust-and-excellence-through-partnership.
- **2025 ADR Awareness Day** – This event further expanded the dialogue, drawing in a wider range of participants and deepening collaboration between European initiatives and stakeholders. Recordings and materials can be accessed here: https://adr-association.eu/events/adr-partnership-awareness-day.

These collaborative ADR Awareness Days not only highlighted the progress of WP3 and the Adra-e project but also strengthened connections within the European ADR ecosystem, laying the groundwork for future initiatives and joint actions.

2.2. Key Impacts following the completion of the tasks

This section summarises the expected impacts of Tasks T3.1 and T3.4 following their completion.

T3.1 Key Impacts

The primary aim of T3.1 is to develop a robust structure and framework for outreach and awareness of ADR, fostering collaboration and alignment among all projects within the Partnership and relevant external initiatives. The expected key impacts include:

- Publication of two open access books capturing the collective knowledge, experiences, and critical outputs of the ADR ecosystem, including the contributions of leading experts and stakeholders. These books serve as enduring educational resources and references for the ADR community across Europe and beyond.
- Establishment of widespread educational and outreach programmes that engage diverse stakeholders and promote the principles of ADR.
- Creation of a dynamic ADR Awareness Centre as an online, living repository of educational resources, best practices, and project results, accessible to all stakeholders.
- Empowerment of European research and innovation activities through the availability and accessibility of resources and materials produced within the ADR ecosystem.
- Enhancement of the EU's attractiveness for talent and expertise in ADR, thus accelerating ADR innovation in Europe.
- Alignment of education and outreach initiatives with the actual needs of stakeholders, ensuring relevance and effectiveness.
- Development of a strong, inclusive, and collaborative network that unites ADR materials and resources from academia, industry, public authorities, and civil society,

including major industrial sectors and all relevant stakeholder groups through the following value propositions:

  - ADR scientists will have access to up-to-date, project-generated resources to support their work.
  - Educators will gain visibility and recognition for their programmes and materials.

- Support for European professionals and organizations in developing collaborations and forming new partnerships.
- The ADR Awareness Centre will accelerate matchmaking processes, helping European organizations and professionals identify project resources and partners for addressing ADR challenges collaboratively.
- Available materials and resources will help stakeholders avoid duplication and maximize the impact of EU investments by promoting synergies and continuous innovation.
- The online repository will serve as a reference point for ADR educational resources during the project and beyond.
- Enhanced information sharing between European projects to foster collaborative learning and mutual growth.
- Facilitation of resource exchange and cross-pollination between Horizon Europe, H2020, and other European initiatives, creating a more synergistic and holistic ecosystem.
- The ADR Awareness Centre will contribute to a unified and efficient ADR ecosystem by sharing project results and relevant initiatives, helping to break silos and encourage cooperative innovation.
- Engagement with the AI-for-Europe project and integration with the AI-on-Demand platform, strengthening Europe's digital innovation landscape.
- Maximization of the impact of EU investments by delivering the ADR Awareness Centre as a valuable, readily available service for stakeholders.
- Dissemination of progress and results, including:
  - The ADR Awareness Day Seminar held on October 20, 2023, via ZOOM (https://adra-e.eu/events/adr-awareness-day-seminar)
  - Collaboration with WP1 on developing a standard outreach message for European projects.
  - Technical collaboration with WP6 on organizing the ADR Awareness Day Seminar.
  - Promotion of the ADR Awareness Centre at the ADR Forum 2023.
  - Publication of two open access books that capture critical outputs and insights from across Europe.

T 3.4 Key Impacts

The primary aim of T3.4 is to promote WP3 and disseminate its outputs through a series of ADR Awareness Day seminars, while also fostering a collaborative culture across the European ADR ecosystem. The expected key impacts include:

- Contribution to open science within the Horizon Europe program by publicly sharing insights, results, and resources.

- Raising awareness of the trustworthiness and acceptability of ADR technologies, targeting key audiences:
    - Public and Citizens
    - Businesses (Large and SMEs)
    - Public Administrators (Large and SMEs)
    - Educators
- Addressing ADR challenges and questions across a large and diverse audience, from insiders (academics and companies) to regional innovation ecosystems and European Member States.
- Facilitating multi-stakeholder dialogue and focused networking through dynamic and engaging event formats.
- Inviting expert speakers and panel members from across Europe to discuss the tasks and outcomes of WP3. This brings cutting-edge knowledge and fosters connections between experts and the audience.
- ADR Awareness Days tailored to different scopes and audiences, connecting the entire ADR ecosystem and encouraging collaboration between diverse actors.
- Increasing awareness and understanding of ADR, helping to overcome one of the most significant obstacles to the uptake of ADR technologies.
- Supporting stakeholders at all stages of the research and innovation cycle:
    - Researchers in public and private institutions
    - Businesses large and small
    - Public sector organizations procuring or adopting ADR technologies
    - Regional and national program leads
    - End-users of ADR solutions
- Generating publishable awareness materials and answers to common ADR questions—potentially in the form of an online “FAQ” hosted on the website—offering enduring and accessible guidance for stakeholders.

3. ADR Awareness Centre

The ADR Awareness Centre is a flagship initiative of the Adra-e project, created to serve as a central hub for education, collaboration, and trust-building in Artificial Intelligence, Data, and Robotics (ADR) across Europe. Rooted in the objectives outlined in the Strategic Research, Innovation and Deployment Agenda (SRIDA), the Centre promotes responsible innovation by making knowledge accessible, supporting stakeholder engagement, and fostering ethical practices in the ADR ecosystem.

This section details the main achievements of T3.1 “ADR Awareness Centre for Education and Outreach” led by UoG in period 2.

3.1. Vision and Objectives of the ADR Awareness Centre

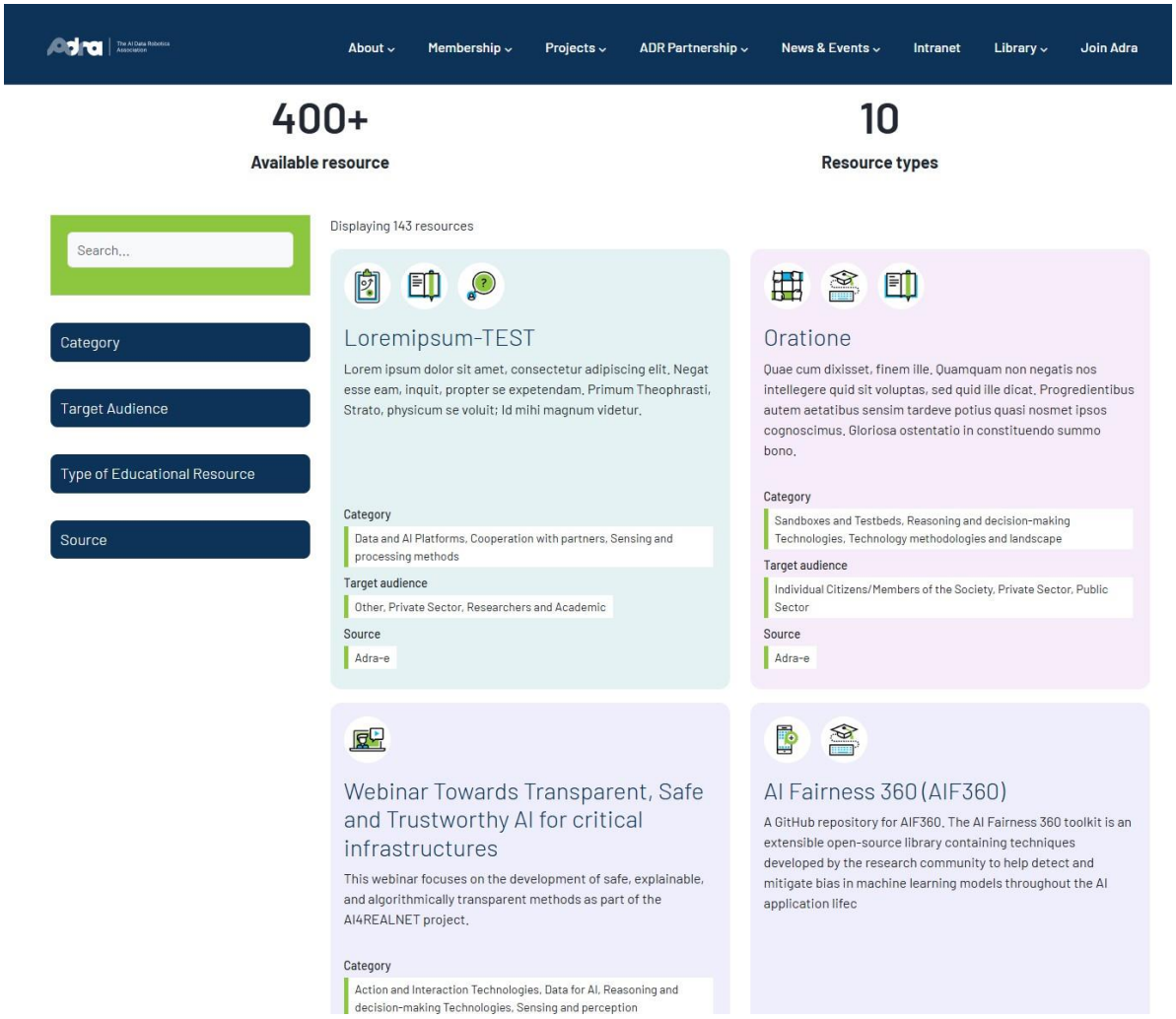
The ADR Awareness Centre was conceived as a strategic initiative to support the responsible advancement of Artificial Intelligence, Data, and Robotics in Europe. At the heart of its mission lies a commitment to empower individuals and institutions through knowledge, trust, and collaboration.

The Centre’s first ambition is to educate a diverse range of stakeholders—industry professionals, researchers, policymakers, and citizens—by providing accessible, high-quality information on the latest developments, trends, and challenges in ADR. In an increasingly complex technological landscape, this educational mission is essential to ensure informed decision-making, responsible innovation, and widespread understanding of ADR’s implications.

Equally central is the goal of promoting trust in ADR systems. The Centre actively highlights the ethical dimensions of technology development, emphasizing transparency, accountability, and societal impact. By doing so, it helps build public confidence and contributes to the broader goal of ensuring that technological progress aligns with European values.

The third pillar of the Centre’s vision is to foster collaboration. By serving as an open and inclusive platform, the Awareness Centre facilitates the sharing of best practices, resources, and innovations among stakeholders across Europe. It brings together perspectives from academia, industry, public institutions, and civil society to create a vibrant, knowledge-driven ecosystem.

Together, these objectives reflect the Adra-e project’s foundational belief: that an informed and connected community is key to the successful and trustworthy integration of ADR technologies into Europe’s digital future.



https://adr-association.eu/adr-awareness-centre

3.2. Overview and key strengths of the ADR Awareness Centre Resources

Below charts provide a clear overview of the distribution of resources within the ADR Awareness Centre, highlighting the focus areas, intended audiences, and formats of the educational materials available.

Category	**Number of assignments** *One resource may include more than one category*
Systems, Methodologies, Hardware, and Tools	261
Data for AI	104
Policy, Regulation, Certificates, and Standards	75
Reasoning and Decision-Making Technologies	29
Action and Interaction Technologies	23
Human-Centric AI	22
Fundamental Rights, Principles, and Values	18
Experiment and Deployment	12
Public Reports	20
Sensing and Perception Technologies	7
Other	6

Target Audience	Number of Assignments *One resource may include more than one category*
Researchers and Academics	95
Professionals	74
Individual Citizens/Members of Society	56
Private Sector	12
Public Sector	12
ADR Experts and Associations	11
Policy Makers	6
Other	4

Type of Educational Resource	Number of Assignments *One resource may include more than one category*
Lectures/Presentations	267
Video Recordings	44
Papers	31
Articles/Books/eBooks	26
Professional Development Resources for Educators	17
Tutorials/How-To Guides	9
Lectures/Presentations with Video Recordings	9
Video/Webinars	8
Software Resources	6
Case Studies	3
White Papers	3
Online Courses	2
Certifications	1
Seminars	1
Datasets	1
Other	11

The ADR Awareness Centre has established itself as a comprehensive, inclusive, and strategically aligned platform supporting the EU's vision for trustworthy and human-centric ADR. The resource statistics reveal several key strengths:

Strong Emphasis on Technical Capacity Building

- With over 260 resources focused on systems, methodologies, hardware, and tools, the Centre is significantly contributing to technical literacy and hands-on support for innovation actors.
- This suggests strong alignment with the needs of developers, SMEs, and researchers engaging in practical ADR development.

Balanced Outreach Across Stakeholders

- The largest audience groups are academics/researchers (43%) and SMEs/start-ups (37%), demonstrating that the Centre is effectively reaching Europe's innovation backbone.
- The inclusion of materials for citizens (18%) and policy makers shows a commitment to democratizing ADR knowledge and ensuring that public trust and governance are not sidelined.

Diverse and Accessible Resource Formats

- The resource library is dominated by lecture-style content (267 resources), backed by videos, guides, books, and professional development resources.
- This diversity supports multiple learning styles, ensures accessibility, and facilitates reuse in educational and training contexts, particularly within academic institutions and lifelong learning initiatives.

Wide Thematic Coverage

- From policy and ethics to data management, human-centric AI, and emerging technologies, the thematic spread indicates that the Centre is not narrowly focused.
- It functions as a one-stop knowledge hub, helping to bridge silos between disciplines, sectors, and user roles in the ADR ecosystem.

Evidence of Strategic Alignment and EU Value Add

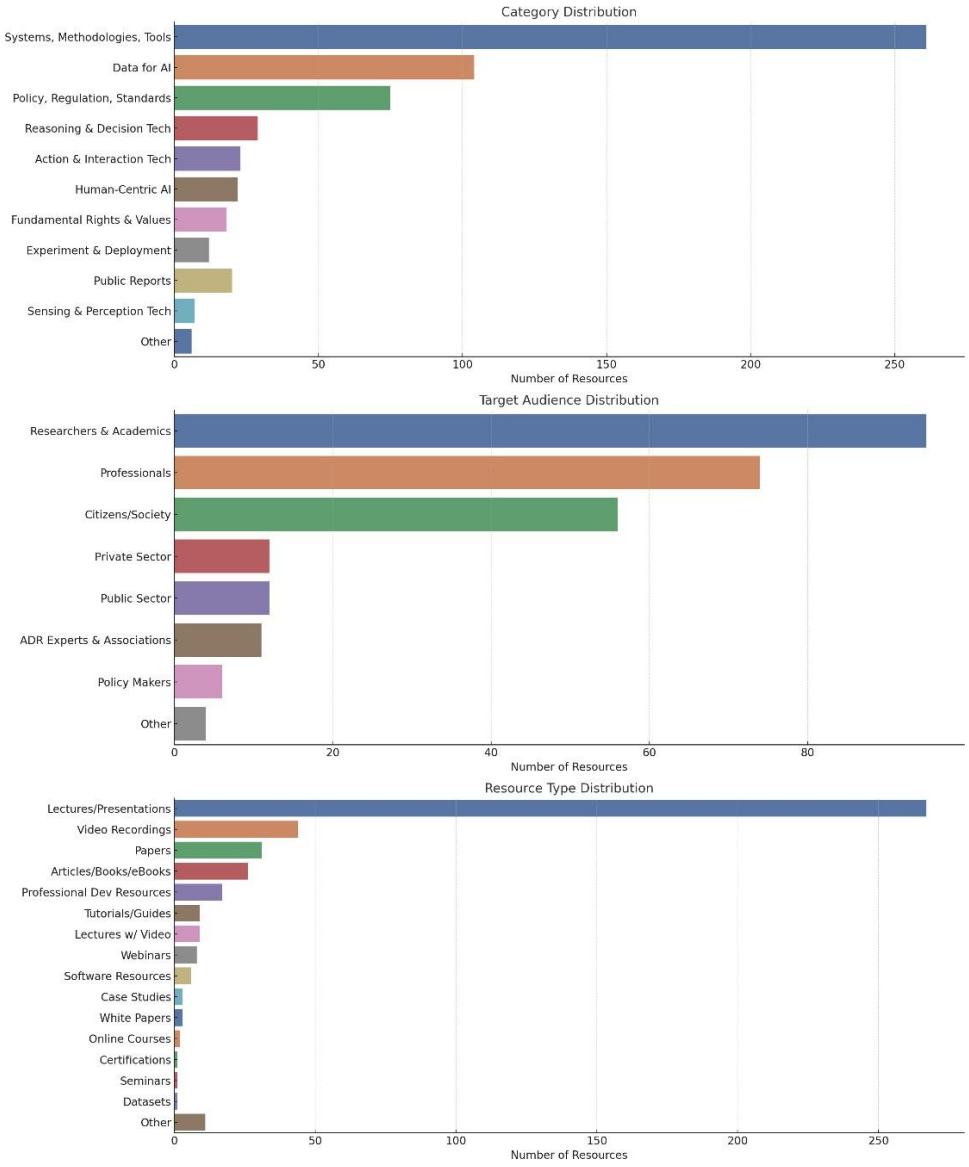
- The inclusion of materials on AI trust labels, regulatory frameworks, and ethical practices reflects a close alignment with the goals of the EU AI Act, Horizon Europe, and digital sovereignty.
- By offering publicly accessible, high-quality materials, the Centre amplifies the value of EU-funded research and promotes cross-project collaboration and transparency.

Summary of the distribution

The ADR Awareness Centre is performing exceptionally well as a flagship knowledge platform. It is:

- Supporting technological advancement through accessible resources,
- Promoting responsible innovation and public trust,
- Reaching diverse European and global audiences, and
- Ensuring that EU ADR efforts are visible, usable, and impactful beyond individual projects.

To maximize future impact, expanding multilingual support, engaging more non-digital sectors, and deepening user analytics could further increase inclusivity and strategic foresight.



3.3. Summary of the Update - ADR Awareness Centre

The ADR Awareness Centre, developed under the Adra-e project, has established itself as a key platform for promoting education, trust, and collaboration in Artificial Intelligence, Data, and Robotics (ADR) across Europe. Since its launch, the Centre has produced over 150 educational resources, engaged more than 3,000 stakeholders, and hosted high-profile events such as the ADR Awareness Day seminars.

The Centre addresses the growing need for public understanding and responsible adoption of ADR technologies by bridging knowledge gaps, promoting ethical practices, and supporting EU priorities for a competitive and inclusive digital economy.

A sustainability plan is in place to ensure the Centre's future impact. It will be fully integrated into the Adra Association**,** ensuring strategic continuity and institutional ownership. Leveraging Adra's infrastructure and continuing collaboration with key projects such as AI4Europe and AI-on-Demand, the Centre is poised to expand its reach, enhance accessibility, and contribute to evidence-based policymaking in the post-Adra-e landscape.

3.4. Sustainability and Ownership

Ensuring the long-term sustainability and continued relevance of the ADR Awareness Centre has been a central consideration from the outset. As the Adra-e project approaches its conclusion, a clear and strategic transition plan has been established to guarantee the Centre's longevity and evolution beyond the project's funded lifecycle.

At the heart of this plan is the integration of the Awareness Centre into the Adra Association—the private side of the European Partnership on AI, Data, and Robotics. This transition will secure institutional ownership and ensure that the Centre remains aligned with the broader strategic priorities of the European ADR ecosystem. As an established actor with deep links to industry, research, and policy communities, Adra is well-positioned to sustain and expand the Centre's role as a central knowledge and engagement hub.

In terms of operations, the Centre will benefit from leveraging Adra's existing digital infrastructure**.** This includes well-established platforms, communication tools, and outreach channels that will be used to maintain the Centre's visibility, update its content, and keep the community actively engaged. The continuity of these tools will allow for a seamless transition, minimising disruption and ensuring that the Centre remains accessible and dynamic.

In parallel, the Centre's collaborative partnerships will continue to play a key role in sustaining content quality and relevance. Ongoing cooperation with projects such as AI4Europe and platforms like AI-on-Demand will enable the exchange of high-quality materials and insights, while deepening interoperability between European initiatives. These collaborations will help ensure that the Centre does not operate in isolation but as part of a broader, synergistic ecosystem focused on responsible and inclusive technological advancement.

Through these measures, the ADR Awareness Centre is well-equipped to thrive as a lasting asset for Europe—continuing to inform, engage, and connect the diverse stakeholders driving the future of AI, Data, and Robotics.

3.5. Rationale for Continuation

As Artificial Intelligence, Data, and Robotics (ADR) technologies become increasingly integrated into vital sectors such as healthcare, manufacturing, mobility, and public services, the need for continuous public awareness and targeted education grows ever more urgent. These technologies, while offering immense potential, also raise complex societal, ethical, and governance challenges that require an informed and engaged stakeholder base.

The ADR Awareness Centre addresses this need directly by serving as a bridge between innovation and society. It works to close knowledge and trust gaps that persist across the ADR landscape, particularly among stakeholders in non-technical, underserved, or less digitally mature sectors. Through its open-access resources and outreach activities, the Centre enables stakeholders to better understand not only the technological dimensions of ADR, but also the ethical, legal, and societal implications.

Beyond its educational function, the Centre supports the European Union's broader policy agenda. It contributes to building a competitive, ethical, and inclusive digital economy—one where technological leadership is underpinned by public trust and equitable access to knowledge. In doing so, it plays a vital role in advancing Europe's digital sovereignty, ensuring that ADR technologies are adopted in a way that reflects shared European values and strengthens democratic resilience across Member States.

The continued operation and evolution of the ADR Awareness Centre is therefore not only desirable but necessary. It is an essential mechanism for fostering responsible innovation, empowering stakeholders, and reinforcing Europe's leadership in human-centric, trustworthy digital transformation.

3.6. Impact to Date

Since its inception, the ADR Awareness Centre has made a measurable and meaningful contribution to strengthening Europe's capacity for responsible innovation in Artificial Intelligence, Data, and Robotics. The Centre has become a vital point of reference for stakeholders seeking to understand, engage with, and contribute to the evolving ADR ecosystem.

A major achievement of the Centre has been the creation of more than 100 educational resources**,** including webinars, articles, policy guides, toolkits, and best-practice materials. These resources are openly accessible and tailored to a diverse audience—from researchers and innovators to policymakers, educators, and the general public. By offering high-quality, well-curated content, the Centre supports continuous learning and empowers stakeholders to make informed decisions about the development and adoption of ADR technologies.

In terms of reach, the Centre has successfully engaged over 3,000 stakeholders from across Europe. These include representatives from academia, industry, public administration, SMEs, and civil society organisations. This broad level of participation illustrates the Centre's ability to cut across sectors and connect communities that are critical to the future of digital transformation in Europe.

Additionally, the Centre has hosted high-profile public events, most notably the ADR Awareness Day seminars**,** which have brought together expert speakers, practitioners, and thought leaders from across the EU. These events have served as a platform for dialogue on urgent topics such as ethical AI, transparency, accountability, and the role of trust in technology. They have also helped to reinforce cross-sectoral collaboration and shared learning, laying the groundwork for continued cooperation across national and institutional boundaries.

Taken together, these outcomes reflect the Centre's growing impact as a trusted, inclusive, and forward-looking initiative—one that is not only disseminating knowledge, but also shaping the way Europe approaches the governance and societal integration of ADR technologies.

3.7. Future Potential

Looking ahead, the ADR Awareness Centre is strategically positioned to deepen its impact and broaden its reach within the evolving European ADR ecosystem. Building on its early successes, the Centre is poised to become not only a hub for knowledge exchange but also a driver of inclusive innovation and policy alignment.

One of the key areas of growth lies in broadening inclusion. The Centre aims to extend its engagement to new audiences that have so far remained on the periphery of ADR discourse—particularly SMEs, startups, and actors in non-digital sectors. These groups are often under-resourced or underrepresented in the adoption of emerging technologies, yet they are critical to ensuring that ADR innovation benefits the whole of society. By reaching these audiences, the Centre can help democratize access to knowledge and empower a more diverse set of contributors to Europe's digital future.

Another priority for the Centre is enhancing accessibility. Recognizing the linguistic and cultural diversity of the European Union, the Centre will focus on producing multilingual content and interactive learning tools tailored to different user needs and skill levels. This will ensure that ADR educational resources are available and understandable to citizens, professionals, and institutions across all Member States—regardless of geography, language, or background.

Finally, the Centre holds considerable potential as a policy-informing platform. With its growing repository of curated resources and its engagement with expert communities, the Centre is well-placed to provide evidence-based insights that can support EU policymakers in shaping forward-looking, human-centric regulations for AI, data, and robotics. By connecting practical experiences with strategic foresight, the Centre can serve as a valuable consultative asset for both national and European-level institutions.

In sum, the ADR Awareness Centre is not only a reflection of what has been achieved, but also a foundation for what is yet to come. Its future lies in expanding inclusion, fostering accessibility, and contributing to a shared policy vision—ensuring that the European ADR ecosystem remains open, ethical, and globally competitive.

3.8. Link with the AI on Demand Platform

A strategic achievement of the ADR Awareness Centre during this reporting period has been the successful integration with the AI-on-Demand (AIoD) Platform, developed under the AI4Europe initiative. The AIoD platform serves as a flagship European infrastructure supporting the development and adoption of trustworthy, human-centric AI, and shares strong synergies with the goals of the Awareness Centre in education, outreach, and community-building.

Following sustained collaboration between the Adra-e and AI4Europe teams, the integration is now fully operational. A dedicated API has been developed, the shared taxonomy has been finalised, and the platforms are now technically and functionally linked. As a result, ADR educational materials and resources produced through Adra-e are now accessible via the AI-on-Demand platform and seamlessly presented under the ADRA Association web platform, creating a unified interface for stakeholders seeking high-quality, curated content on AI, Data, and Robotics.

This integration delivers several key benefits:

- Content Interoperability: Resources published through the Awareness Centre are now discoverable within the AIoD ecosystem, enabling greater reuse and visibility.
- Increased Reach and Impact: The joint presentation under the ADRA Association web platform ensures broader dissemination to both technical and non-technical audiences across Europe.
- Operational Efficiency: The use of a shared taxonomy and API streamlines content management and avoids duplication, maximizing the value of EU-funded efforts.
- Platform Synergy: The integration reinforces the collaborative positioning of both platforms, offering users a more coherent, enriched, and accessible experience when navigating Europe’s digital ecosystem for trustworthy AI.

This milestone marks a significant step toward the sustainability and long-term scalability of the ADR Awareness Centre. It demonstrates how European partnerships can create

interoperable, federated digital infrastructures that amplify impact while supporting the European Commission's goals for openness, trust, and coordination in the digital domain.

Going forward, the ADRA Association will continue to maintain and evolve this integration as part of its broader mission to support an inclusive, transparent, and trustworthy ADR ecosystem in Europe.

3.9. Key Performance Indicators (KPIs) and Monitoring Framework

To ensure that the ADR Awareness Centre meets its objectives and delivers measurable value to the European ADR ecosystem, a structured KPI and Monitoring Framework was established as part of WP3. This framework serves both as a management tool and an accountability mechanism, enabling the Adra-e team to track progress, assess outcomes, and make evidence-based adjustments where needed. The KPIs span quantitative and qualitative metrics aligned with core goals of the Centre, including educational resource development, stakeholder engagement, platform interoperability, outreach impact, and long-term sustainability. Regular monitoring activities—combined with stakeholder feedback, web analytics, and milestone reviews—have been used to evaluate performance throughout Period 2 (M18–M36) and to guide the Centre's continued evolution beyond the Adra-e project lifecycle.

Objective	*Indicator*	*Target / Baseline*	*Actual (Period 2: M18–M36)*	*Status / Notes*
1. Build and maintain a central online repository for ADR education and outreach	No. of educational resources created and published	100 resources	150+ resources	Target exceeded; new resources are submitted;
	Platform functionality and uptime	Functional, >95% uptime	Achieved	Fully operational; hosted under ADRA web platform
2. Expand stakeholder engagement across Europe	No. of engaged stakeholders across academia, industry, public sector	2,000 users	Over 3,000 (including visiting users and event participants)	Target exceeded; ongoing monitoring via analytics
	Geographical diversity of users	Representation from 10+ countries	Participants from 15+ EU and associated countries	Strong engagement across regions
	No. of organizations contributing/sharing content	15 contributing orgs	24 organizations (private and public)	Broadening engagement
3. Foster visibility and interoperability	Integration with AI-on-Demand platform	Conceptual design in M18	Full technical integration with API and shared taxonomy (M34)	Achieved and operational under ADRA

with European platforms				Association platform
	No. of shared resources between platforms	N/A (baseline: 0)	200+ resources integrated/shared	Alignment complete; content synchronised
4. Promote awareness of ethical and trustworthy ADR	No. of ADR Awareness Day events organized	3 events (2023, 2024, 2025)	3 completed (with recordings, panels, audience feedback)	Well-attended; recordings archived and promoted via Centre
	No. of participants in awareness events	250 (cumulative target)	500+ participants	Target significantly exceeded
	Diversity of speakers and topics	Experts from academia, policy, industry	Achieved (panelists from across sectors and regions)	Demonstrated cross-sectoral relevance
5. Ensure long-term sustainability and transition	Ownership and maintenance plan approved	By M36	Transition to Adra Association completed (incl. web hosting)	Fully integrated and governed post-project
	Continued collaboration agreements (e.g., AI4Europe)	1–2 formalised collaborations	Achieved: AI4Europe, AI-on-Demand, and additional project links	Ensures ongoing content flow and alignment with EC priorities

During Period 2 (M18–M36), the ADR Awareness Centre demonstrated strong performance across all key performance indicators (KPIs), significantly exceeding expectations in several core areas.

A major success was the development and publication of over 100 educational resources, including webinars, guides, articles, and toolkits. These materials have been carefully curated to meet the needs of different stakeholder groups and are now fully accessible through the ADR Awareness Centre, which remains stable, active, and hosted under the ADRA Association's digital platform.

The Centre also achieved extensive stakeholder engagement, attracting over 3,000 users from across academia, industry, government, SMEs, and civil society. Participation came from more than 20 EU and associated countries, reflecting both the Centre's pan-European relevance and its success in reaching diverse audiences. In total, 22 organizations actively contributed content, including Adra-e consortium members, educators, and external collaborators.

A key milestone in this period was the successful technical integration with the AI-on-Demand platform, enabled by the finalization of a shared taxonomy and API. This integration ensures seamless content exchange and interoperability between platforms and positions the

Awareness Centre as a core node in Europe's digital infrastructure for trustworthy AI, Data, and Robotics.

The ADR Awareness Day events, held in 2024 and 2025, were also significant achievements. These events attracted over 750 participants and brought together expert speakers and panelists from across different sectors and domains to discuss topics such as ethical AI, transparency, and responsible innovation. The Centre's ability to convene such diverse voices demonstrates its value as a space for meaningful dialogue and knowledge exchange.

Crucially, the Centre's sustainability has been secured through its formal transition to the Adra Association, which now oversees its continued operation, integration, and alignment with broader ADR objectives. Strategic collaborations with AI4Europe and the AI-on-Demand platform will continue to support content flow and ensure long-term relevance.

Overall, the ADR Awareness Centre has not only met but exceeded its deliverable targets. It has established itself as a trusted, inclusive, and impactful platform that supports the European ADR ecosystem through education, visibility, and collaboration.

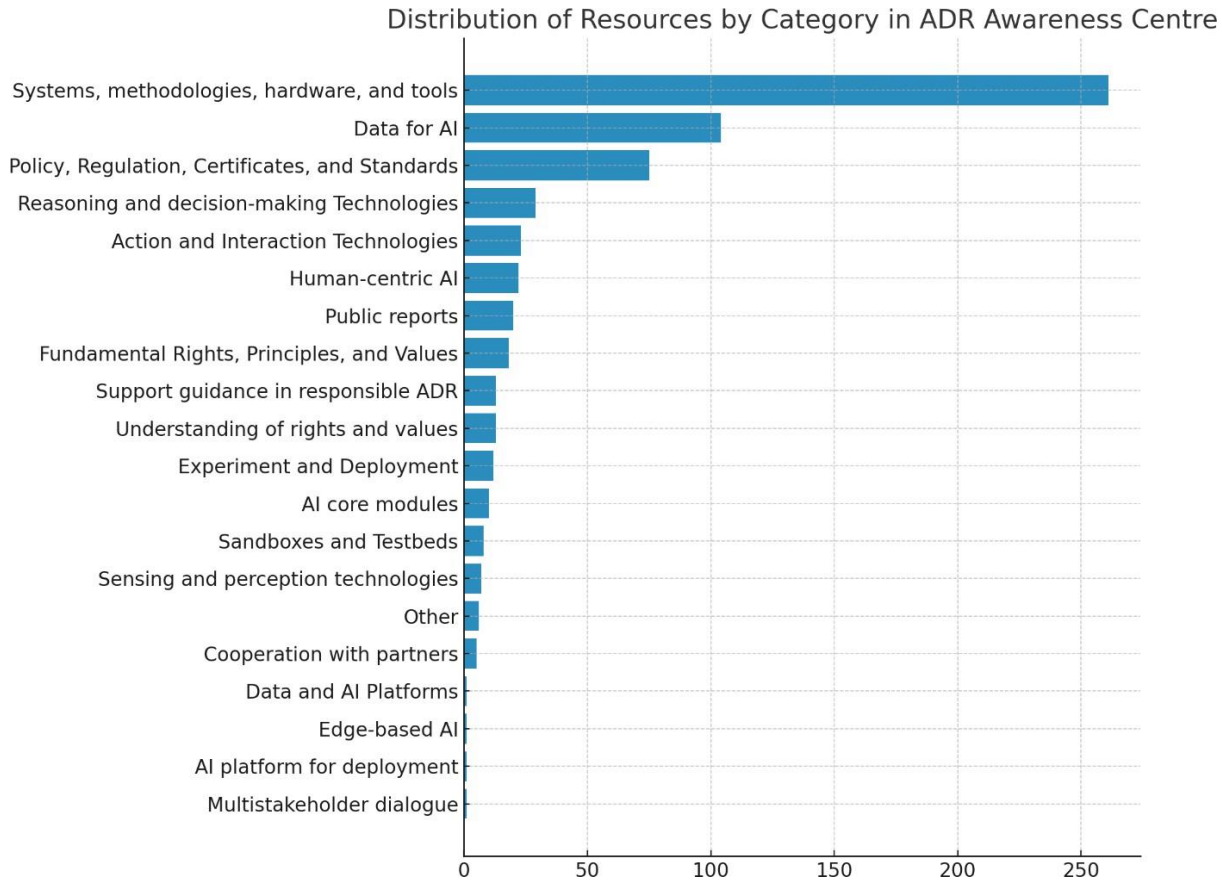
In summary, the ADR Awareness Centre has established itself as a vital resource for Europe's digital future. It offers more than just information—it creates a dynamic, participatory space for shaping how ADR technologies are developed, governed, and trusted. With its transition to the Adra Association, the Centre is set to continue advancing a responsible, human-centric, and forward-looking ADR agenda for Europe.

3.10. SRIDA-Aligned Resource Analysis and Strategic Relevance of Awareness Centre Submissions

Now that the ADR Awareness Centre's taxonomy has been explicitly designed based on the Strategic Research, Innovation and Deployment Agenda (SRIDA) framework, it offers a significant advantage over other existing ADR repositories. Unlike generic collections, this SRIDA-aligned structure allows for systematic classification, policy-relevant tracking, and strategic insight into the thematic focus of available resources. It provides a clear lens through which to assess how well current submissions reflect European priorities for trustworthy and value-driven AI, Data, and Robotics. This section presents an initial analysis of the distribution of submitted resources across SRIDA-defined categories—highlighting strengths, gaps, and potential areas for further development.

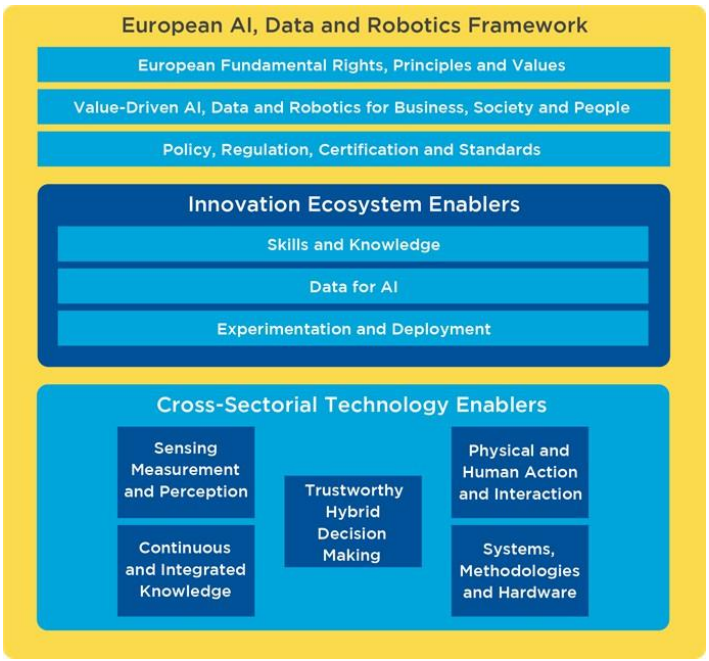
Below is the bar chart showing the distribution of resources by category *(one resource may include more than one category)* in the ADR Awareness Centre. It clearly highlights that:

- Systems, methodologies, hardware, and tools is by far the most populated category.
- Data for AI and Policy, Regulation, Certificates, and Standards follow.
- Reasoning technologies also hold significant weight.
- Several specialized categories like Edge-based AI or Multistakeholder dialogue have very few entries, indicating potential gaps or niche focus.



The SRIDA framework is structured around three layers:

1. European AI, Data, and Robotics Framework (top-level principles and governance)
2. Innovation Ecosystem Enablers (skills, data, experimentation)
3. Cross-Sectorial Technology Enablers (technical capabilities)



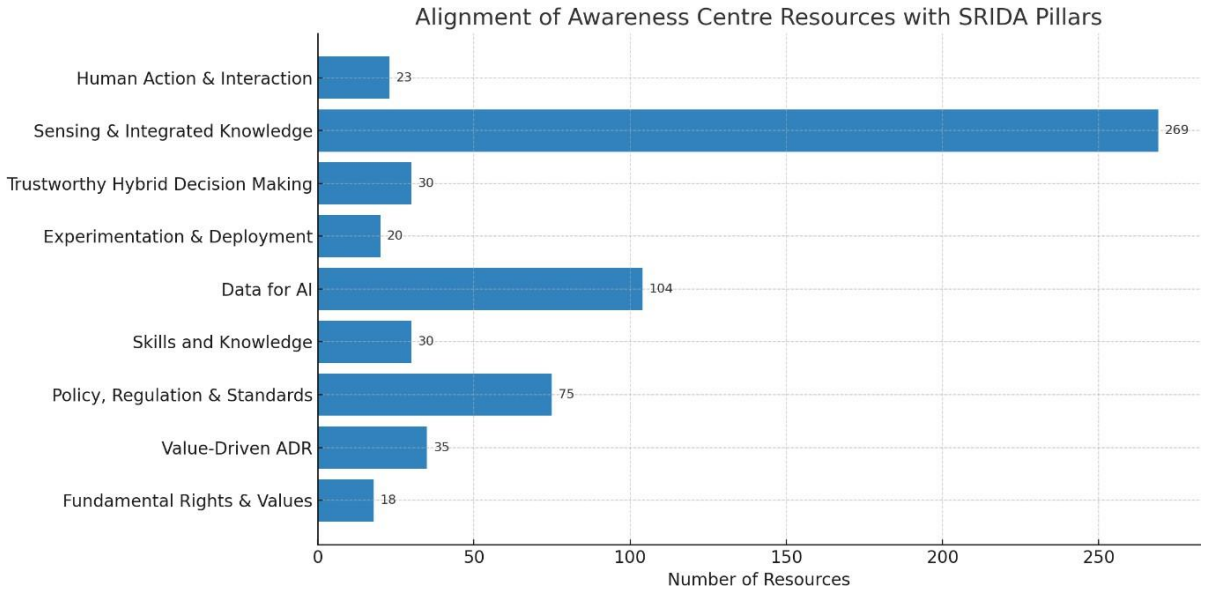
We mapped these layers against the Awareness Centre resource submissions:

SRIDA Component	Matching Awareness Centre Category	# Resources
Fundamental Rights & Values	Fundamental Rights, Principles, and Values	18
Value-Driven ADR	Human-Centric AI; Support for Responsible Implementation	35 (22 + 13)
Policy, Regulation, Certification	Policy, Regulation, Certificates, and Standards	75
Skills and Knowledge	Public Reports (policy/educational use); AI Core Modules	30 (20 + 10)
Data for AI	Data for AI	104
Experimentation and Deployment	Experiment and Deployment; Sandboxes and Testbeds	20 (12 + 8)
Trustworthy Hybrid Decision Making	Reasoning and Decision Technologies; Trustworthiness	30 (29 + 1)
Sensing, Perception, Integration	Sensing and Perception; Systems and Tools; Edge-based AI	269 (7 + 261 + 1)
Human Action and Interaction	Action and Interaction Technologies; Robotics	23 + (part of Systems)

Below is the SRIDA alignment matrix showing how ADR Awareness Centre resources are distributed across the key pillars of the SRIDA framework:

- Sensing & Integrated Knowledge and Data for AI are the most well-represented areas.
- Categories such as Policy & Regulation and Trustworthy Decision-Making are moderately supported.
- Fundamental Rights, Skills, and Experimentation & Deployment remain underrepresented.

The below chart offers a strategic lens into where awareness efforts are concentrated—and where more targeted contributions may be needed to support a balanced, SRIDA-aligned ecosystem.



Key Insights

- Overwhelming emphasis is placed on Systems, Methodologies, and Tools (261 submissions), suggesting the community's priority is still heavily technical and infrastructure-oriented.
- Policy, certification, and governance resources (75) show good alignment with SRIDA's top layer, reflecting growing awareness of responsible AI.
- Data for AI is well-covered (104), matching the central role of data access and quality in innovation enablement.
- Skills, experimentation, and trust are underrepresented (e.g., 30 for Skills/Knowledge, only 20 for Deployment), which are vital for real-world adoption and scaling.
- Topics related to ethics, human interaction, and societal values are gaining traction (35 total), but require further amplification to match the ambitions of value-driven ADR in SRIDA.

What Can Be Learned

- Technical domains dominate ADR awareness materials—especially tools, systems, and data handling.
- Responsible AI and trust mechanisms are recognized, but not as widely documented or diversified across domains.
- The ADR Awareness Centre reflects solid alignment with SRIDA but shows imbalances in thematic focus, particularly in experiential learning, skill-building, and applied deployment insights.

WP3 Recommendations

- Rebalance Resource Development - Invest in generating or curating more materials focused on:
  - ✓ Experimentation and real-world deployment
  - ✓ Cross-sectoral skills and knowledge sharing
  - ✓ Human-in-the-loop systems and inclusive design

- Promote Interlinkage Across Categories - Use tagging or interactive navigation to help users connect, for example, "Data for AI" with “Experimentation” or “Fundamental Rights” with “Decision Making.”
- Amplify Value-Driven Narratives - Encourage submissions and content that frame ADR innovations around societal values, inclusion, and resilience, reinforcing the top level of the SRIDA agenda.
- Guide Contributors with SRIDA Tags - Adapt the taxonomy submission template to align directly with SRIDA categories to promote strategic completeness.
- Monitor Balance with KPIs - Update your Awareness Centre’s monitoring framework to track thematic balance aligned with SRIDA pillars (e.g., share of “deployment” vs. development”).

4. ADR Awareness Days

As part of Task 3.4 under WP3, the Adra-e project organized two high-impact ADR Awareness Day events in 2024 and 2025**,** both delivered in a fully virtual format. These events were conceived as flagship outreach and engagement activities to promote the goals of the ADR Awareness Centre, raise public awareness of trustworthy and ethical ADR practices, and showcase key outputs of the Adra-e project.

The virtual format allowed for broad and inclusive participation, attracting a diverse audience from across Europe, including representatives from academia, industry, public administration, SMEs, and civil society. Each event featured expert panels, keynote presentations, and interactive discussions on critical themes such as transparency in AI, societal trust, and the role of education and certification in shaping responsible ADR ecosystems.

Designed to reflect the Centre's multi-stakeholder vision, the ADR Awareness Days served as both a dissemination platform and a dialogue forum, creating opportunities for cross-sector exchange and strengthening connections within the wider ADR community. The 2024 and 2025 editions built upon one another, amplifying impact, expanding outreach, and reinforcing the Centre's role as a trusted source of knowledge and collaboration in the European ADR landscape.

4.1. Objectives

The primary objective of Task 3.4 is to raise awareness and foster stakeholder engagement around the acceptability, trustworthiness, and societal impact of AI, Data, and Robotics (ADR). Through a series of high-visibility annual events, this task aims to:

- Promote key outputs from WP3, including the ADR Awareness Centre (T3.1), the meta-analysis of ADR-related skills and externalities (T3.2), and the AI Trust Label and its methodology (T3.3).
- Create space for multi-stakeholder dialogue that includes perspectives from public administrators, businesses (both SMEs and large enterprises), educators, NGOs, and the general public.
- Enhance public understanding of ethical ADR practices and support the wider adoption of trustworthy technologies.
- Amplify the impact of Adra-e by providing a platform for showcasing outcomes, sharing best practices, and aligning with EU digital policy goals.

These events serve as a strategic dissemination tool, helping to extend the reach and influence of the project beyond its immediate network.

4.2. Implementation plan

The 2024 and 2025 ADR Awareness Day events were carefully designed to maximize accessibility, impact, and environmental sustainability. While initial plans considered co-locating or co-timing the events with the ADR Forum and the Convergence Summit, a strategic decision was made to host both events virtually. This approach ensured greater inclusivity across Europe, simplified logistics, reduced travel-related emissions, and enabled easier participation from a wide range of stakeholders.

- The 2024 Awareness Day (held in December) focused on "Building Trust and Excellence Through Partnership," was an online event aimed at fostering dialogue and collaboration

on pivotal topics influencing the future of AI, Data, and Robotics in Europe. The event focused on the convergence of AI and Data regulation, interoperability, trustworthy AI, generative AI, and foundational models, with a particular emphasis on their integration into robotics technologies.

- The 2025 ADR Awareness Day aimed to foster multi-stakeholder engagement and showcase leading-edge applications of ADR in critical sectors such as healthcare, manufacturing, and human–machine collaboration. Hosted by the WP3 lead organization, the objective was to promote trust, ethics, and societal value in the deployment of ADR technologies, and to support dialogue between researchers, industry, and policy stakeholders across Europe.

Both events are structured to include keynote presentations, expert panels, and interactive Q&A sessions, with an emphasis on cross-sector collaboration, policy relevance, and practical application of trustworthy ADR principles. The virtual format has allowed the project to reach a pan-European audience, facilitate deeper engagement through online platforms, and maintain cost-efficiency while delivering high-quality content.

4.3. ADR Awareness Day 2024

Date: 17 December 2024
Time: 10:00–12:00 CET
Format: Virtual
Location: Online
Registration Link: Zoom Webinar Registration
Recording: Watch the recording
Presentations: Download the presentations

Purpose:

The 2024 ADR Awareness Day, themed "Building Trust and Excellence Through Partnership," was an online event aimed at fostering dialogue and collaboration on pivotal topics influencing the future of AI, Data, and Robotics in Europe. The event focused on the convergence of AI and Data regulation, interoperability, trustworthy AI, generative AI, and foundational models, with a particular emphasis on their integration into robotics technologies

Key Themes:

Towards an ecosystem of trust in AI, Data and Robotics introduces trust and accountability across the ever-changing scenes of artificial intelligence, data and robotics. This session will look at the state of play in Europe as we progress in the creation of an ecosystem where stakeholders-from developers to end-users-can engage in the exploration of Trustworthy AI, data and robotics. Join us as we dive into building a trusted ecosystem that fosters transparency, encourages value-driven practices, and supports sustainable innovation.

Towards an Ecosystem of Excellence in AI, Data and Robotics focuses on European Excellence and best practices in AI, Data, and Robotics that leverages the transformational enablers of Generative AI and Data Spaces. This session will explore the benefits which arise from these enabling technologies being integrated into enhancing the performance of AI, data sharing, and robotics systems through improved efficiency, effectiveness, and quality. The highlights will be provided on how to achieve European excellence in these fields with respect to our mission, which is to ensure high-level standards and operational excellence across the ADR landscape.

Agenda:

- 10:00–10:10: Objectives and Introduction to the ADR Partnership and ADR
  Speaker: Philip Piatkiewicz, Secretary General - AI, Data and Robotics Association
- 10:10–10:20: ADR Ecosystem Supports
  Speaker: Jozef Geurts, Deputy Director European Partnerships Department, Inria
- 10:20–10:50: Achievements of the Partnership (ADRA)
  Speaker: Emanuela Girardi, President - AI, Data and Robotics Association
- 10:50–11:30: Panel Discussion
  Panelists:
  - Kimmo Rossi, Head of Sector - DG CONNECT, European Commission
  - Cem Gulec, Policy and Programme Officer - DG CONNECT
- 11:30–11:50: Panel Discussion - Towards Ecosystem of Trust in ADR
  Panelists:
  - Stefan Leijnen, Head of EU Relations - AiNed
  - Fabio Pianesi, Distinguished Fellow - FBK and Senior Advisor for EU Affairs - ARUBA/Dynamo
  - Nabil Belbachir, Director BoD - euRobotics aisbl
  - Robert Bruckmeier, General Manager, Computing and Artificial Intelligence Network - BMW Group
- 11:50–11:55: Call for Book Chapters
  Speaker: Heike Vornhagen, Insight Centre for Data Analytics
- 11:55–12:00: Closing Remarks
  Speaker: Philip Piatkiewicz, Secretary General - AI, Data and Robotics Association

Specific input from WP3:

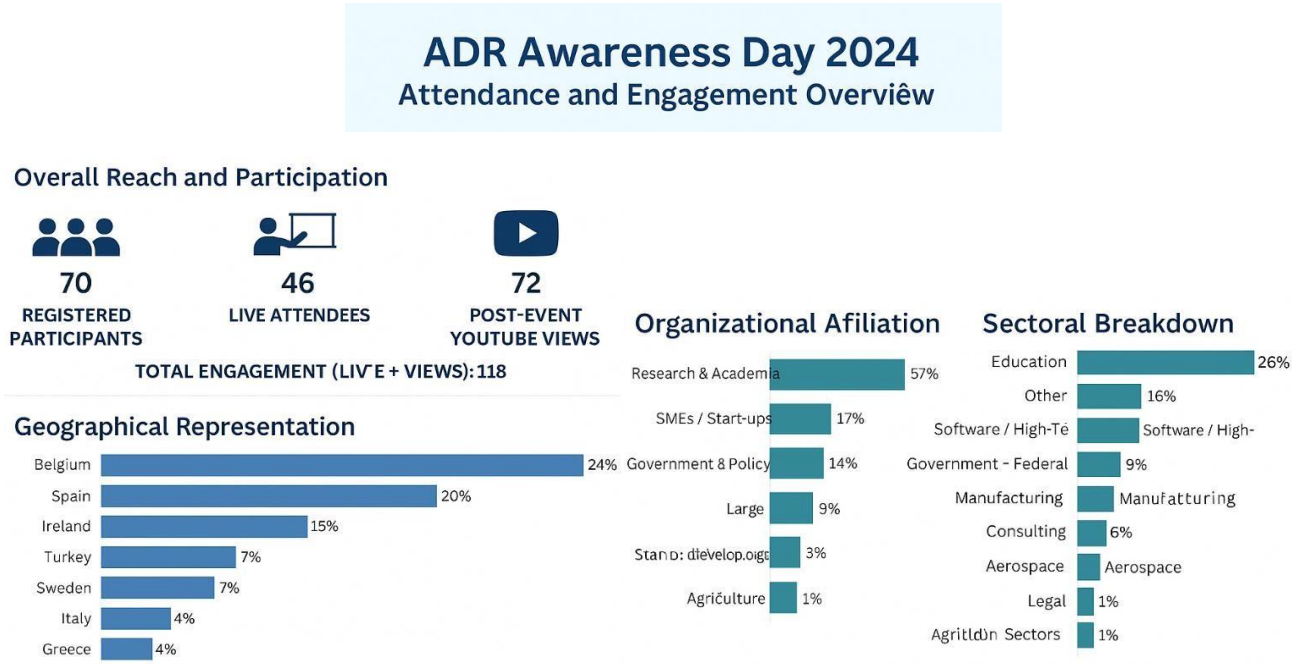
As part of its contribution to the 2024 ADR Awareness Day, Work Package 3 (WP3)—led by the University of Galway—played a key role in shaping the substance and direction of the event. In particular, WP3 proposed a set of challenge questions designed to guide the panel discussions and stimulate meaningful dialogue among experts from policy, industry, academia, and civil society. These questions were closely aligned with WP3's focus on trust, education, and responsible ADR adoption, and helped ensure that the event's content was grounded in the realities and priorities of diverse stakeholder groups.

To promote transparency, accessibility, and knowledge sharing, the full recording of the event has been made publicly available. Stakeholders who were unable to attend the live session can now access the complete recording online via the following link: Watch the 2024 ADR Awareness Day Recording

This ensures continued visibility and engagement with the topics explored during the event, extending its reach and relevance well beyond the day of broadcast.

Outcome of the event:

The event successfully brought together policymakers, researchers, innovators, industry leaders, and stakeholders from the ADR ecosystem to discuss and collaborate on building a trusted and excellent European ADR landscape. The discussions and insights shared during the event are expected to inform future initiatives and strategies within the ADR community.



Attendance and Engagement Overview:

Overall Reach and Participation

- Registered Participants: 70
- Live Attendees: 46
- Post-Event YouTube Views: 72

- Total Engagement (Live + Views): 118

The event reached a broader audience beyond the live session through the publicly available recording, nearly doubling its visibility post-event.

Geographical Representation

Participants came from a wide range of EU countries and one key non-EU country, reflecting strong pan-European interest and engagement.

- Top Countries Represented:
  - Belgium: 24%
  - Spain: 20%
  - Ireland: 15%
  - Turkey (Non-EU): 11%
  - Sweden: 7%
- Additional participation from: Italy, Greece, France, Portugal, Poland, Netherlands, Luxembourg, and Germany.

Note: EU countries made up the majority of attendees, demonstrating the regional relevance of the ADR topics discussed.

Organizational Affiliation

Attendees represented a diverse set of organizations, with the largest segment from academia and research institutions:

- Research & Academia: 57%
- SMEs / Start-ups: 17%
- Government & Policy: 14%
- Large Enterprises: 9%
- Standard Development Orgs & Professional Associations: 2% combined

This composition illustrates strong academic engagement alongside growing interest from innovators and public-sector actors.

Sectoral Breakdown

Participants brought perspectives from various sectors, with particular representation from education, high tech, and manufacturing:

- Education: 26%
- Software / High-Tech: 20% combined
- Government – Federal: 9%
- Manufacturing: 7%
- Consulting & Aerospace: 9% combined
- Additional sectors included Legal, Transportation, Energy, Agriculture, and Accounting (each at 1%).

Insight: The distribution confirms the cross-sectoral relevance of trustworthy AI, data, and robotics, especially in education, innovation, and public service.

This detailed attendee profile reflects the success of the virtual format in reaching a geographically and professionally diverse audience, further validating the decision to hold the event online for greater accessibility and reduced environmental impact.

Agenda and Structure:

The seminar consists of an introduction and overview of the ADR ecosystem, a keynote and two panel discussions.

Session 1: Introduction, overview and keynote

Session moderator: Philip Piatkiewicz

ADR ecosystem support by Joost Geurts

Joost Geurts, from Inria's European partnership department, presented an overview of the ADRA project, which supports the AI, Data and Robotics (ADR) partnership. Now in its final phase, ADRA aims to build a trustworthy and excellent European ADR ecosystem. Geurts highlighted the foundational influence of the European Commission's 2020 white paper, emphasizing that trust and excellence are interlinked and essential for innovation and economic growth. The partnership, a contract between the Commission and ADRA representing the private sector, is built around mobilising the community to actively engage with shared objectives. Since July 2022, the Adra-e project has supported this effort by fostering collaboration, sustainability, and coherence across AI, data, and robotics stakeholders, with a strong emphasis on community building at its core.

Several key initiatives have emerged from the project. A Strategic Roadmap, developed through ongoing community engagement, guides research and innovation priorities. An Awareness Center has been created to offer educational resources via the AI-on-Demand platform, helping stakeholders better understand the ADR landscape. ADRA has also organized major events like ADR Awareness Day and the ADR Forum in Eindhoven to promote community interaction and visibility. Upcoming initiatives such as the Convergence Summit and the Future Ready Initiative are designed to strengthen collaboration and ensure the long-term sustainability of Europe's ADR ecosystem.

Keynote: Achievements of the Partnership (ADRA)

Speaker: Emanuela Girardi, President of Adra

The speaker shared ADRA's progress and vision, emphasizing the need for Europe to act decisively in AI, data, and robotics to remain globally competitive. A video from a U.S. factory illustrated how advanced robotics, powered by AI and data, could revolutionize industry—an opportunity Europe must seize. ADRA, established in 2021 by five key European associations, now represents 160 members across 26 countries. Its main objectives include developing the Strategic Research and Innovation Agenda (SRIDA), fostering a cohesive ecosystem, advancing standardisation, and promoting European leadership in generative AI and robotics. Central to ADRA's approach is engaging a wide range of stakeholders—from industry leaders and startups to researchers and policymakers—ensuring their active participation in shaping Europe's technological future.

ADRA has made significant strides in uniting the AI, data, and robotics communities, supporting strategic initiatives, and producing key deliverables such as SRIDA. Their 15 topic groups and sector-specific task forces address vital challenges and promote responsible, energy-efficient AI technologies aligned with European values. With priorities like achieving full-stack AI sovereignty, accelerating AI Factories, and investing in European generative AI champions, ADRA plays a

crucial role in Europe's digital and green transformation. Its efforts in standardisation and regulatory guidance aim to empower even smaller players to navigate new legislation. As the ecosystem continues to grow, ADRA encourages more stakeholders to join, underscoring the importance of collective action in shaping a trustworthy and competitive European AI landscape.

Session 2: Panel 1 - European Commission's Role in AI, Data, and Robotics

Session moderator: Prof Edward Curry

Panel speakers:

Kimmo Rossi, Head of Sector - DG CONNECT, European Commission

Cem Gulec, Policy and Programme Officer - DG CONNECT

Challenge questions for Panel 1:

1. Why is it important for Europe to have a partnership on ADR?
2. With the rapid development of Generative AI (GenAI), how can Europe ensure these tools align with the principles of trustworthy AI, particularly in data privacy and ethical design?
3. What are data spaces, and how do they enable secure and scalable data sharing across industries/sectors?
4. How can Europe leverage initiatives like Data Spaces and AI Factories to establish global leadership in AI, Data, and Robotics (ADR)?
5. What measures are needed for a cohesive European ecosystem that drives excellence in ADR?
6. What support will the AI office provide to the ADR ecosystem?

Summary of the discussion:

The European Commission has launched the ADR partnership to unite efforts across AI, data, and robotics by aligning public and private investment (€1.3 billion each). Central to this effort is Adra, a hub fostering collaboration and guiding funding through initiatives like SRIDA and Horizon Europe. The Commission also introduced the AI Act, the world's first comprehensive AI law, which uses a risk-based approach and is implemented by the newly formed AI Office. This office supports compliance, supervises General Purpose AI, and funds trustworthy AI research—recently investing €65 million into projects on explainable and robust AI systems.

To overcome barriers to data sharing, the Commission is rolling out Common European Data Spaces—secure, interoperable infrastructures designed to enable safe exchange of industrial and sensitive data. These spaces, along with newly launched AI Factories, aim to support the entire AI value chain, reduce complexity for innovators, and foster trust and compliance. Moving forward, Europe's strategy focuses on investment, coordinated regulation, talent retention, and global collaboration to position itself as a leader in ethical and competitive AI development.

Session 2: Panel 2 - Towards ecosystem of trust in ADR

Session moderator: Philip Piatkiewicz

Panel speakers:

Stefan Leijnen, Head of EU relations – AiNed

Fabio Pianesi, Distinguished Fellow - FBK and Senior Advisor for EU Affairs - ARUBA/Dynamo

Nabil Belbachir, Director BoD - euRobotics aisbl

Robert Bruckmeier, General manager, Computing and Artificial Intelligence Network - BMW Group

Challenge questions for Panel 2:

1. How do we ensure trustworthiness across the entire lifecycle of ADR technologies—from design to deployment and decommissioning?
2. How can we bridge the gap between public perception and the reality of ADR systems' capabilities and limitations?
3. How do we address ethical concerns about data ownership and consent in ADR applications?
4. What role do international standards organizations play in advancing trust and reliability across ADR technologies?
5. What emerging trends or technologies are likely to shape the landscape of trust in ADR over the next decade?
6. How has ADR Partnership addressed these challenges and what will happen next?
7. What are the main obstacles your industries face in scaling AI and robotics solutions, whether in terms of infrastructure, workforce readiness, or regulatory constraints?
8. How do you see current and upcoming regulations—such as the EU AI Act—shaping the adoption and development of AI and robotics in industry? Are there specific areas where more clarity is needed?

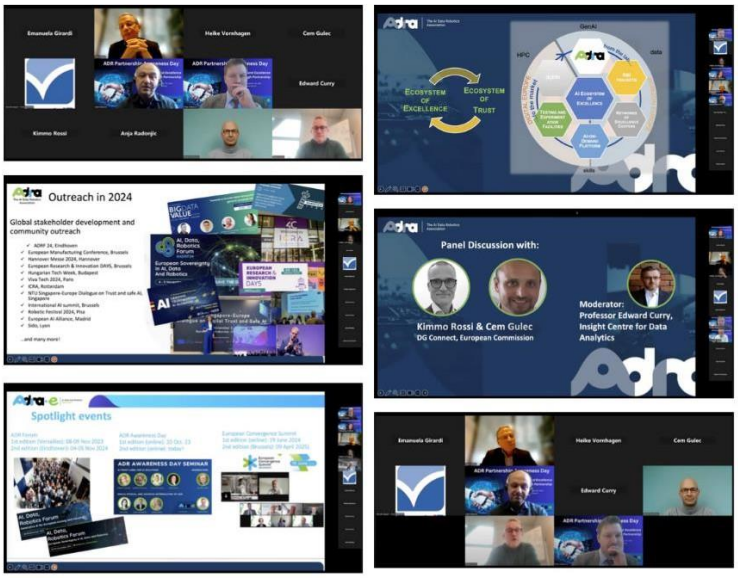
Summary of the discussion:

Europe's ability to compete globally in AI, data, and robotics hinges on its strengths in data management, industrial sectors, and emerging technologies. Stefan points out that while AI dominance has been concentrated in American companies, Europe can capitalize on innovations in data collection, efficient algorithms, and next-generation hardware, such as neuromorphic computing and energy-efficient edge computing. This opportunity arises as we approach "peak data," requiring a shift in how AI is scaled, giving Europe a chance to lead.

Robert emphasizes that AI, data, and robotics are too important to be left solely in the hands of a few dominant companies, especially from the U.S. He sees "peak data" as an opportunity for Europe, where it excels in industrial data management. This shift in the AI landscape presents an opening for Europe to strengthen its technological position, but it requires active investment in scaling and deploying technologies, not just research.

Nabil stresses the interconnected nature of AI, data, and robotics, noting that Europe has strong robotics companies but must also strengthen its position in AI, especially in areas where these domains converge. He emphasizes the need for investment in both technological advancements and talent retention to stay competitive in the global market. Without attracting and retaining skilled individuals, even significant investments may not yield the desired outcomes.

Fabio highlights the crucial role of standards in positioning Europe for success. He explains that standards are easier to implement and adapt than complex regulations, making it easier for businesses to scale solutions across borders. By integrating regulatory requirements into standards, Europe can create a more favourable environment for AI development, fostering global competitiveness while ensuring alignment with ethical and legal frameworks.



4.4. ADR Awareness Day 2025

Date and time of third Awareness Day: 27th May 2025 14:00 – 16:00 CEST

Location: Virtual event via ZOOM

General description of the event (advertised):

The ADR Awareness Day is an annual stakeholder engagement event dedicated to raising awareness of the acceptability and trustworthiness of ADR (AI, Data, and Robotics) technologies. Hosted by the Adra-e Project, the seminar brings together experts and stakeholders to explore the latest developments and initiatives aimed at fostering trust and ethics in AI applications. The ADR Awareness Day Seminar is a flagship event designed to spark meaningful dialogue and showcase leading-edge projects focused on the ethical, responsible, and impactful use of AI, Data, and Robotics (ADR). As part of the ongoing “Building Trust and Excellence Through Partnerships” initiative, this year's seminar brings together thought leaders, researchers, and innovators to explore how ADR technologies are shaping the future of industry, healthcare, and human-machine collaboration.

This year’s theme focuses on real-world applications and pilot projects that are not only technically groundbreaking but also grounded in societal values and ethical considerations. The seminar will spotlight how ADR is enabling transformation across key sectors, with a special focus on trustworthiness, human-centric design, and sustainability.

What to expect

- FP10 and the Future of ADR: Join a forward-looking session on the European Commission’s recently published proposal for the 10th Framework Programme for Research and Innovation (FP10). This discussion will explore how ADR priorities can align with FP10’s strategic objectives, offering insights into upcoming opportunities for collaboration and funding.
- Companion Robots in Elderly Care: Explore a novel robot platform designed to investigate the technical challenges and ethical dimensions of robotic companionship in caregiving. This session will address the delicate balance between autonomy, empathy, and safety.
- AI for Smarter Production: Discover innovative AI-based management approaches that are reshaping production systems, enhancing decision-making, and boosting operational efficiency in modern manufacturing environments.

- Industry 5.0 and Sustainability: Dive into new applications and pilot cases where ADR technologies support the shift toward a more sustainable, circular manufacturing ecosystem, aligned with the principles of Industry 5.0.
- Human-in-the-Loop & Explainable AI: Learn how combining human oversight with explainable AI techniques enhances trust and transparency—particularly in high-stakes or complex decision-making environments.
- Next-Generation Human-Robot Interaction: Get a glimpse of the future of industrial collaboration, where seamless human-robot teamwork drives productivity, safety, and adaptability in evolving workspaces.

Whether you're an industry expert, researcher, policymaker, or curious stakeholder, this seminar is your gateway to the partnerships, projects, and perspectives shaping the future of trustworthy ADR.

Target audience

This seminar is intended for professionals, researchers, policymakers, and anyone interested in the ethical and societal implications of AI, Data, and Robotics technologies. It offers a valuable opportunity to engage with experts in the field, gain knowledge about ongoing projects, and participate in discussions that shape the future of ADR.

Agenda

Time	Topic	Speaker
14:00	Welcome and introduction to the Awareness Day	Philip Pietkiewicz - Adra
14.05	Introduction to ADRA Introduction to and Overview of the new FP10	
14.25	Q&A	
Presentations		
14.35	AI-based Management Approaches for Production Processes	Robert Woitsch - BOC Group
14.45	Data, AI and Robotics transformative power in Industry 5.0: new applications and pilot cases for a more Sustainable and Circular Manufacturing.	Sergio Gusmeroli - Politecnico di Milano
14.55	A Companion Robot Platform for Exploring Technical and Ethical Aspects in Elderly Care	Lorenzo Boi - University of Macerata
15.05	The cross- fertilization between the Human-in-the-Loop Approach and Explainable AI techniques towards Trustworthiness	Theodore Dalamagas - "Athena" Research and Innovation Center
15.15	ARISE: Advancing Industrial Collaboration - The Next Generation of Human-Robot Interaction	Mireya de Diego Moro - CARTIF

15.25	Q&A	Amaya Garmendia - Adra Heike Vornhagen - Insight Centre for Data Analytics
15.45	Close	Amaya Garmendia - Adra

Speakers

Philip Pietkiewicz
Adra
Philip Piatkiewicz is a seasoned leader in European affairs and project management, currently serving as the head of Adra – the AI, Data, and Robotics Association. Adra is a leading strategic technology network of stakeholders from academia, industry, and the public sector, dedicated to advancing and adopting AI, data, and robotics technologies across Europe. Adra acts as the private side for the AI, Data, Robotics public-private partnership in Horizon Europe. Philip's expertise encompasses strategic planning, stakeholder engagement, policy analysis, and project execution. He has successfully managed and coordinated complex collaborative projects that support the research, innovation, and industrial agendas of regional, national and EU level, with a particular focus on key enabling technologies and digital industries.

Sergio Gusmeroli
Politecnico di Milano
Sergio Gusmeroli is Research Coordinator at Politecnico di Milano, Industrial Engineering Department. He has been and is coordinating several H2020 and HEP projects in the domain of Digital Manufacturing, especially focusing on Data Technologies and Artificial Intelligence. Sergio is also co-leading the BDVA working group about Smart Manufacturing Industry and is work package leader of the SM4RTENANCE deployment action in the Digital Europe Programme about Data Spaces for Manufacturing.
The key question to be debated in the panel is: how is the convergence of AI Data and Robotics solutions affecting sustainability and resilience of Manufacturing value chains?

Lorenzo Boi
University of Macerata
Lorenzo Boi's is currently pursuing a PhD in Teaching and Learning Sciences at the University of Macerata. His research explores human-robot interaction, artificial intelligence, and blockchain, aiming to bridge the gap between advanced digital systems and real-world applications.

Mireya de Diego Moro
CARTIF
MSc. Eng. in Industrial Electronics and Automation and Degree in Telecommunications Engineering from the University of Valladolid, Spain. Since 2013, she has been a dedicated researcher at CARTIF's Industrial and Digital Systems Division, specializing in robotics, automation and process control within the agri-food manufacturing industry. Her expertise spans across multiple EU projects, where she has driven innovation and sustainability across diverse industrial sectors. Currently, she serves as the coordinator for the ARISE project, focusing on advancing human-robot interaction (HRI) technologies in manufacturing, logistics, and healthcare. Her work aims to enhance efficiency, safety and collaboration between humans and robots in various applications, reflecting her commitment to pushing the boundaries of industrial automation and robotics.

Robert Woitsch
BOC Group
Dr. Robert Woitsch is member of the board at BOC Products & Services AG and responsible for Research & Innovation. He is involved in innovation projects, mainly funded by the European Commission, addressing knowledge-, information- and data management, Industry 4.0 as well as

software engineering and cloud computing since 2000. His project portfolio contains typically 3 to 6 EU-Projects at a time. Currently, he coordinates the artificial intelligence and robotic project FAIRWork. Robert acted as expert for the European Commission as well as a speaker and reviewer in scientific conferences. He is also active in customer consultancy projects, in the domain of knowledge-, quality- and risk-management. The research results are distributed in the ADOxx development community, a world-wide acting Open Innovation Community with over 7.000 developers collaborating in the domain of conceptual modelling. Recently he is involved in building industrial laboratories at BOC in the context of the OMiLAB community.

Theodore Dalamagas
"Athena" Research and Innovation Center
Theodore Dalamagas is Research Director and Vice Director of Information Management Systems Institute at ATHENA Research Center, as well as co-founder and Research Director of Symbiolabs (spinoff of ATHENA). He received his Diploma in Electrical Engineering from NTU Athens, Greece (1995), his MSc in Advanced Information Systems from Glasgow University, Scotland (1997), and his PhD from NTUA (2004).
He has more than 20 years of R&I experience of running and coordinating EU and national R&I IT projects. His research and technology areas of interest include: scientific databanks and e-research infrastructures, data Web and information retrieval, data interoperability and integration, as well as data services and data science applications for several domains (e.g., health and genomics, industrial biotechnology, manufacturing). He has published more than 100 articles in international journals and conference proceedings.

Specific input from WP3

- UoG proposed speakers to showcase projects in the area of manufacturing, healthcare and human-in-the-loop.
- UoG contacted all speakers, coordinated presentations and facilitated the Q&A session.
- For people unable to attend the seminar, a complete recording of the session is available at https://www.youtube.com/watch?v=vydGgF8iqxU&t=3s.
- Presentations from the event are available at https://adr-association.eu/events/adr-partnership-awareness-day.

Outcome of the event

The 2025 ADR Awareness Day successfully engaged over 70 registrants, with 46 live participants and 72 YouTube views, resulting in a total outreach of 118 participants and viewers. Attendees represented a wide cross-section of Europe, spanning 13 countries and sectors such as academia (57%), SMEs (17%), government and policy (14%), and large enterprises (9%).

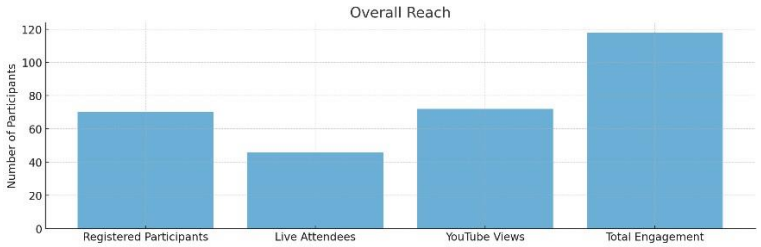
The event's presentations and panel sessions sparked active engagement from the audience. Participants posted a series of insightful questions, leading to dynamic discussions around the following key areas:

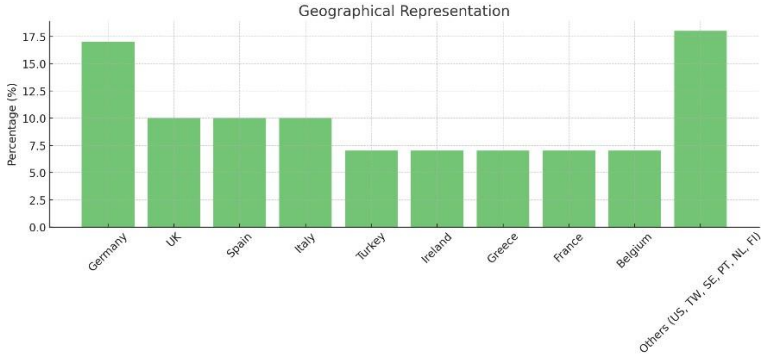
- Trustworthy AI and robot certification: How assurance mechanisms and trust labels can be adapted to robotics technologies.
- Bias and human-in-the-loop systems: While intended to reduce historical bias, the inclusion of human oversight also raises questions about new forms of bias introduced through judgment inconsistencies.
- User testing challenges: Issues such as the psychological vulnerabilities of test users and the feasibility of simulating real-world diversity of scenarios (e.g., multiple user types, environments, and edge cases) were debated.
- Replicability of lessons learned: A nuanced discussion unfolded around which project insights are generalizable across domains, and which are highly context-dependent.

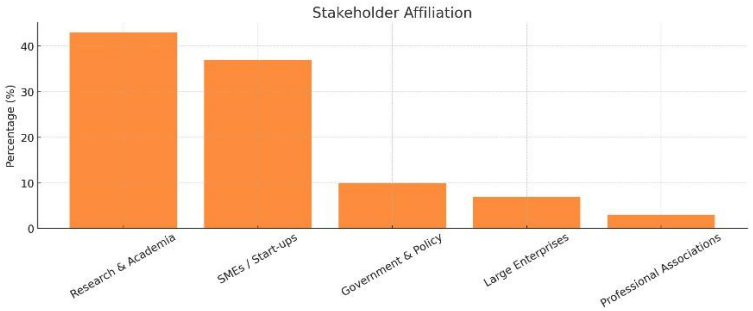
These discussions reflected the high level of expertise and interest among participants and reinforced the importance of multi-stakeholder dialogue in shaping trustworthy ADR ecosystems.

Moreover, the event provided a critical platform to showcase the outputs of WP3, particularly in relation to the ADR Awareness Centre, the meta-analysis of ADR externalities, and the development of AI trust labels. By grounding technical innovation in societal values and cross-sectoral engagement, the event made a tangible contribution to Adra-e's mission of promoting ethical, inclusive, and impactful ADR development in Europe.

The full event recording remains publicly available, enabling continued visibility and engagement beyond the live session: . Watch the recording







Statistics on Attendees

Overall Reach

- Registered Participants: 70
- Live Attendees: 46
- YouTube Views (Post-Event): 72
- Total Engagement (Live + Views): 118

The event's virtual format enabled broad accessibility and post-event reach, nearly doubling live attendance through recorded viewership.

Geographical Representation

Participants joined from both EU and non-EU countries, highlighting the global relevance of the themes discussed. The top countries represented include:

- Germany (17%)
- United Kingdom, Spain, and Italy (10% each)
- Turkey, Ireland, Greece, France, and Belgium (7% each)
- Additional participation from: United States, Taiwan, Sweden, Portugal, Netherlands, and Finland (3% each)

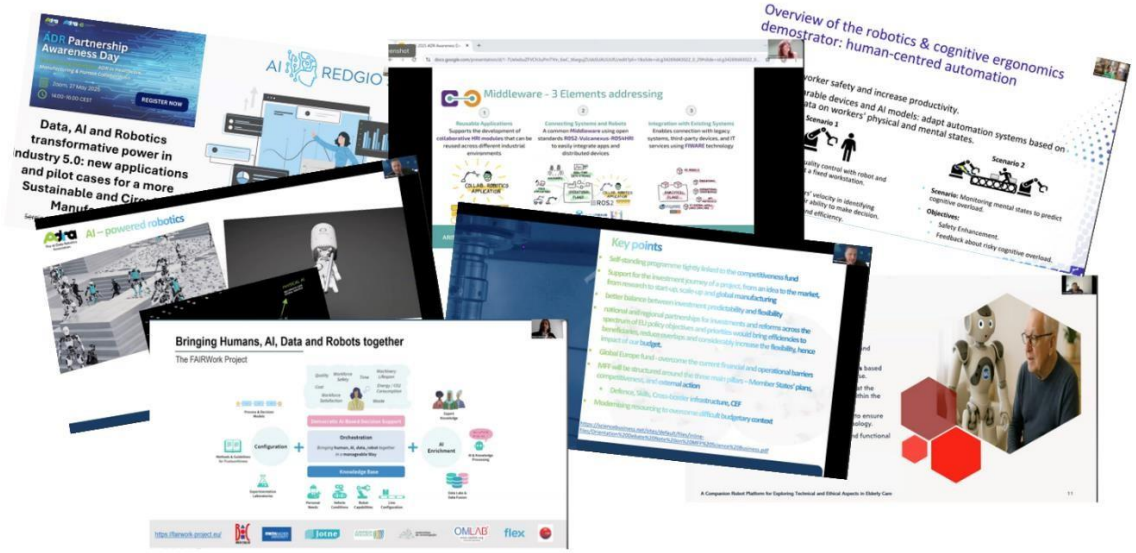
Participants came from 16 countries, with EU member states making up the majority, but with notable engagement from global actors in the UK, US, and Asia.

Stakeholder Affiliation

Attendees came from a variety of institutional settings:

- Research & Academia: 43%
- SMEs / Start-ups: 37%
- Government & Policy: 10%
- Large Enterprises: 7%
- Professional Associations: 3%

The strong presence of researchers and SMEs underscores the event's dual emphasis on scientific rigor and applied innovation, while participation from policymakers and industry reflects cross-sectoral interest.



5. Dissemination of the ADR Content: Open Access Books

5.1. Book 1. Artificial Intelligence, Data and Robotics: Foundations, Transformations and Future Directions

Call For Chapters

Artificial Intelligence, Data and Robotics: Foundations, Transformations, and Future Directions

Editors

Edward Curry, University of Galway, Ireland.
Fredrik Heintz, Linköping University, Sweden.
Heike Vornhagen, University of Galway, Ireland.
Ahmed Nabil Belbachir, NORCE Norwegian Research Centre, Norway.
Philip Piatkiewicz, Adra, Belgium.
Emanuela Girardi, Pop AI, Italy.
Marc Schoenauer, INRIA, France.
Juha Roning, University of Oulu, Finland.

Overview

Artificial intelligence, Data and Robotics (ADR) present an opportunity and a challenge for our society, a chance to improve the competitiveness of the public and private sectors, and a challenge to translate core ADR technical strengths to deliver industrial competitiveness, societal well-being, and environmental sustainability.

The European Partnership on AI, Data, and Robotics has mobilised the ADR ecosystem in Europe to provide strong leadership in these areas, both in science, innovation, and deployment. It creates dialogues that address fundamental issues around deployment and citizen trust in AI. It enables a rich ADR innovation ecosystem built on Europe's many strong components, from its strong academic excellence, strong skills pipeline, and global companies to its innovation-driving regulation and standards coupled with best practices. The vision is to lead the world in researching, developing, and deploying value-driven, trustworthy ADR based on fundamental European rights, principles, and values.

This Open Access book will study the foundations of the emerging ADR research and innovation ecosystem and the associated enablers necessary to deliver value-driven ADR for business and society. The book is a compilation of selected high-quality chapters covering best practices, technologies, experiences, and practical recommendations on research and innovation.

Aims and Goals

The aim of the book is to educate the reader on techniques and technologies for AI, Data and Robotics. The book will explore cutting-edge theories, technologies, methodologies, and best practices within industrial and public sector scenarios. The book provides the reader with a basis for understanding the scientific foundations, the transformation potential of the technologies, and the future research challenges. The book will have two primary audiences. First, researchers and students in the fields of AI, Data and Robotics and associated disciplines, e.g. computer science, information technology, and information systems among others. Second, industrial practitioners will find practical recommendations based on rigorous studies that contain insights and guidance

on ADR across several technology and innovation areas. Finally, the book would interest local, national, and international policymakers.

Topics

Topics of interest to the book include, but are not limited to, the following:

Area A: Foundations

- Edge AI
- Distributed AI
- Hybrid AI Models and Architectures
- Human-in-the-loop
- Data-driven AI Models
- Cognitive Computing
- Sensing and Perception
- Knowledge and Learning
- Action and Interaction
- Neuromorphic AI
- AI and Data Life Cycle Management and Governance
- Standards for AI and Data Interoperability
- Data Enrichment for AI and Robotics
- Reasoning and Decision-Making
- Systems, Methodologies, Hardware and Tools.
- Standardisation and Interoperability
- Novel design concepts for flexible robot end-effectors
- Human-robot collaboration

Area B: Transformations

- AI and Data for Robotics
- Autonomous Robots
- Grand Challenges
- Transformation in Manufacturing
- Transformation in Healthcare
- Transformation in Telecommunications
- Transformation in Industrial Optimization
- Transformation in Cognitive Ergonomics

Area C: Future Directions

- Forward-looking roadmaps for research and innovation
- Open challenges

Table of Contents

Area A – Foundations

- Shaping Europe's Future: The Convergence of AI, Data, and Robotics
  Edward Curry, Fredrik Heintz, Heike Vornhagen, Ahmed Nabil Belbachir, Philip Pietkiewicz, Emanuela Girardi, Marc Schoenauer, Juha Roning.
- The cross-fertilization between the Human-in-the-Loop Approach and the Explainable AI Techniques towards Trustworthiness

Marina Da Bormida in Cugurra, Sotiris Koussouris, Daniele Crippa, Carl Hans, Robert Hellbach, Andre Tabone, Dimitris Bibikas

- Hybrid Intelligence: The fusion of science-based and Machine Learning models
  Eleni Lavasa, Theodora Chrysoula, Vasileios Gkolemis, Charalambos Sinnis, Theodore Dalamagas, Sotiris Koussouris, Nefeli Bountouni, Konstantinos Perakis, Viktor Daropoulos, Charalambos Lambri, George Ioannou, George Pallis, Marios Dikaiakos, Daniele Crippa, Andre Tabone, Dimitrios Bimpikas, Stratos Keranidis, Carl Hans, Robert Hellbach, Dimitris Bouras
- Frugal Machine Learning for Energy-efficient, and Resource-aware Artificial Intelligence
  John Violos, Konstantina-Christina Diamanti, Ioannis Kompatsiaris, and Symeon Papadopoulos
- SRIDA: Charting the Future of a Progressive, Inclusive, and Sustainable European ADR Ecosystem
  Fredrik Heintz, Katerina Linden
- Generative artificial intelligence to tackle visual data accessibility challenges
  Lore Goetschalckx, Kaili Wang, Siri Willems, Tom De Schepper

B – Transformations

- The role of Digital Transformation in Manufacturing under the light of AI integration
  John Angelopoulos, Kosmas Alexopoulos
- AI-based Management Approaches for Production Processes – Identifying Requirements for Frameworks and Architectures
  Anna Sumereder, Robert Woitsch, Damiano Falcioni, Marlene Mayr
- Data, AI, Robotics transformative power in Industry 5.0. New applications and pilot cases for a more Sustainable and Circular Manufacturing
  Gusmeroli Sergio, Dalle Carbonare Davide, Anibal Reñones, Gema Antequera Garcia, Carmen Polcaro, Cinzia Rubattino, Gabriella Monteleone, Pietro Pittaro, Juanan Arreta, Marielena Marquez Barreiro, Santiago Muños-Landin, Lacalle Ignacio, Jesus Alonso
- Towards the Irish Mobility Data Space: Challenges, Opportunities, and Requirements for Data Sharing
  Rafiqul Haque, Diarmuid Ó Conchubhair, Muhammad Asif Razzak, Majjed Al-Qatf, Fatemeh Ahmadi Zeleti, Wassim Derguech, Edward Curry
- Towards a Holistic Framework for Human-AI Collaboration in Safety-Critical Systems
  Milad Leyli-abadi, Ricardo J. Bessa, Daniel Boos, Clark Borst, Alberto Castagna, Ricardo Chavarriaga, Duarte Dias, Adrian Egli, Andrina Eisenegger, Joost Ellerbroek, Anna Fedorova, Cristina Felix, Anton Fuxjäger, Joaquim Geraldes, Samira Hamouche, Mohamed Hassouna, Sjoerd Kop, Bruno Lemetayer, Giulia Leto, Roman Liessner, Jonas Lundberg, Antoine Marot, Maroua Meddeb, Manuel Meyer, Hélio Sales, Viola Schiaffonati, Manuel Schneider, Irene Sturm, Julia Usher, Herke Van Hoof, Jan Viebahn, Toni Waefler, Mouadh Yagoubi & Giacomo Zanotti
- CyclOps: Leveraging Semantic Technologies for AI and Data Life Cycle Management and Governance
  Monica Caballero, Anna Queralt, Diego Calvanese, Alexandros Oikonomidis, Chukwuemeka Muonagor, Alberto Abella, Piero Campalani, Benjamin Cogrel, Gabriele Tassi, Julia Palma, Alex Barceló, Saeed Talebzadeh, Ross Little, Peio Oiz, Florencia Pérez, Alexandros Kalafatelis, Maria Font, Sergi Nadal, Alexandros Nizamis, Admela Jukan, Simone Martin Marotta, Ricardo Simón-Carbajo, Javier Conejero, Dolores Ordóñez

- Intelligent underwater perception: current trends and future directions
  Nicoletta Risi, Gabriel Guimarães Carvalho, Olivier Vasseur, Phil Reiter, Tom De Schepper
- Brain-to-Speech: Prosody Feature Engineering and Transformer-Based Reconstruction
  Mohammed Salah Al-Radhi, Géza Németh, Andon Tchechmedjiev, Binbin Xu
- A Companion Robot Platform for Exploring Technical and Ethical Aspects in Elderly Care
  Lorenzo Boi, Slivia M Massa, Diego Reforgiato Recupero, Daniele Riboni, Rubén Alonso, Michele Cardinali, Elena Ricci, and Alberto Pirni
- Robotics in Healthcare IoT: and Deep Learning Perspectives
  A'adel Sayyahi, Seyed Enayatallah Alavi, Morteza Jaderyan

C – Future Directions

- Roadmaps and Agendas for Research and Innovation in Artificial Intelligence, Data, and Robotics
  Paulo Figueiras, Giorgos Ioannou, Charalambos Lambri, Sangheeta Reji, Elena Mossali, Martina Imarisio Neviani, Sotiris Koussouris, Nefeli Bountouni, Marina Da Bormida in Cugurra, Robert Hellbach, Dimitris Bibikas, Philip O'Brien, Carlos Agostinho, Ricardo Jardim-Gonçalves
- Advancing Industrial Collaboration: The Next Generation of Human-Robot Interaction
  Anibal Reñones, Mireya de Diego, Francisco Melendez, Zoi Arkouli, Nikos Dimitropoulos, Christos Gkrizis, Fernando Castaño, Rodolfo E. Haber and Leire Bastida
- Human-AI Interaction and Visualization Perspectives on ADR
  Kostiantyn Kucher, Magnus Bång, Jonas Lundberg

5.2. Book 2. Trustworthy Generative AI: Fundamentals, Applications, and Future Directions

Call For Chapters

Editors

Edward Curry, University of Galway, Ireland.
Fredrik Heintz, Linköping University, Sweden.
Stefan Leijnen, HU University of Applied Sciences, Netherlands.
Thibauld Jongen, Common Sense Robotics, Belgium.
Fatemeh Ahmadi Zeleti, University of Galway, Ireland.
Ahmed Nabil Belbachir, NORCE Norwegian Research Centre, Norway.
Philip Piatkiewicz, Adra, Belgium.
Emanuela Girardi, Pop AI, Italy.

Overview

Artificial Intelligence is reshaping industries, economies, and daily life. The rise of Generative AI powered by developments in large language models and natural language understanding has created new opportunities in content creation, autonomous systems, and interactive dialogues. This book aims to provide a comprehensive edited volume that brings together contributions from leading experts and practitioners in Trustworthy Generative AI. It is a compilation of selected high-quality chapters covering best practices, technologies, experiences, and practical recommendations on research and innovation. It delves into the foundational principles, cutting-

edge techniques, and diverse applications of Trustworthy Generative AI, and systematically investigates theoretical foundations with practical insights.

This open-access book will capture the key concepts, strengths, and limitations required to build Generative AI systems that can solve complex problems and uphold transparency, fairness and human-centric values to make them trustworthy. It gives the reader a comprehensive interpretation of emerging technologies in generative AI, exploring their potential while examining the ethical implications of their use.

Aims and Goals

The book aims to provide readers with a thorough understanding of Trustworthy Generative AI, focusing on the foundations required to build Trustworthy Generative AI systems that drive innovation. It explores generative models, natural language understanding, and conversational AI advancements. Through pragmatic frameworks, real-world applications, and case studies, the book equips AI practitioners with the tools to implement Trustworthy Generative AI solutions. Each chapter is authored by a distinguished researcher or practitioner, giving readers a rich and diverse perspective on the state-of-the-art. The book serves as a reference guide for researchers and professionals seeking to deepen their understanding of this rapidly evolving field.

Topics

Topics of interest to the book include, but are not limited to, the following:

Fundamentals:

- NeuroSymbolic AI
- Conversational AI and Agents
- Data for Trustworthy Generative AI, including data preprocessing, data spaces, and provenance.
- Large Language Models/Hybrid Language Models
- Natural Language Processing and Understanding
- Multimodal & Multilingual
- Training and Optimization Techniques
- Evolution of generative models
- Content creation
- Autonomous Systems
- Generative AI for Robotics
- Trust, Ethics, and Regulation, including Ethical and Societal Implications, including Bias and fairness
- Regulatory and legal issues, including copyright and privacy concerns

Applications

- Industry Case Studies and Real-world Applications of Trustworthy Generative AI
- Advancing Healthcare, including medical imaging, drug discovery, & personalised medicine
- AI for Science and Education
- Media, Tourism, Culture, Arts, and Creative Industries
- Manufacturing, Automotive, Engineering, and Industrial Sectors, including Generative AI for Robotics
- Societal Challenges including the Environment, Climate Change, Democracy, Migration, Energy, Mobility, and Ageing Populations.
- Government and Public Sector
- Legal, Finance, and Insurance

Fundamentals

- Data Compliance Services: A Key to Unlocking Data Governance Under the AI Act
  Ivo Emanuilov, Dessislava Petrova-Antonova, Sylvia Ilieva
- Trustworthy Generative AI in Practice: Automating EU AI Act Compliance Assessment
  Nataša Krčo, Ioanna Angeliki Kapetanidou, Srdjan Krčo, Alexandros Nizamis, Anja Jakovljević, Konstantinos Votis
- Building Trust in Synthetic Media: Provenance Frameworks for Generative AI
  Mansura Habiba, Declan Joyce
- Mitigating Bias and Enhancing Trust in AI-based Systems through a Synthetic Data-driven Framework
  Dimitrios I. Fotiadis
- The Integrity Challenge of AIaaS Model
  Georgios Spathoulas
- Generating Bias? Auditing Fairness in Synthetic Tabular Data from Generative AI
  Swati Swati, Arjun Roy, Emmanouil Panagiotou, Eirini Ntoutsi
- Opening the Black Box: Explaining Generative AI Processes with Scene2Model
  Wilfrid Utz, Iulia Vaidian, et al.
- Responsible Development of Generative AI for Societal and Economic Resilience
  Michael McCann, Muslim Jameel Syed
- AI Bill of Materials (AI BOM) for Trustworthy Generative AI
  Mayuko Kaneko
- Cognitive Load, Safety, and Generative Co-Pilots: Insights from Engineering and Neuropsychology
  Mario Espinosa Gámez
- Generative AI and the Erosion of Cognitive Autonomy: A Neuroethical Framework
  Mario Espinosa Gámez

Applications

- Responsible AI in Practice: Designing a Trustworthy GenAI Chatbot in Public Administration
  Gizem Gezici, Chandana Sree Mala, Fosca Giannotti, Sezer Kutluk
- Trustworthy Generative AI for Unmanned Vehicle Swarms
  Mai Le, Barry O'Sullivan, Hoang D. Nguyen
- From Bench to Bedside: Auditing Trustworthiness of Generative AI in Healthcare Practice
  Alin Navas
- Trustworthy Generative AI for Real-Time Travel Risk Management: Personalised Safety at Scale
  Dr Sheelagh Brady
- Advancing User-Centric Manufacturing in High-Impact Industrial Sectors through Multimodal GenAI
  Sotiris Makris

- Bridging the Trust Gap: Hybrid Decision Support Systems as a Pathway to Trustworthy Generative AI
  Artur Bogucki
- Generative Models in Medical Imaging: when Less is More
  Khaoula ElBedoui, Walid Barhoumi
- Exploring Diagnostic Innovations: Survey of Deep Learning Models for Lung Disease Classification
  Hoda El-Batrawy, Wesam Ahmed
- Manufacturing, Automotive, Engineering, and Industrial Sectors, including Generative AI for Robotics
  Mahmood Seyyedzadeh
- Generative AI for Robotics
  Mahmood Seyyedzadeh, Amir A. Ghavifekr

6. Conclusion

This deliverable has presented a comprehensive account of the progress and implementation of Task 3.1 (ADR Awareness Centre) and Task 3.4 (ADR Awareness Days) during Period 2 (M18–M36) of the Adra-e project. These two tasks represent the cornerstone of WP3's outreach, engagement, and capacity-building objectives, serving to bridge gaps between innovation, education, and trust in AI, Data, and Robotics (ADR).

The ADR Awareness Centre, developed under T3.1, has become a dynamic and evolving EU-wide platform for knowledge dissemination. With over 100 curated resources, cross-project integration via API with the AI-on-Demand platform, and sustained engagement from over 3,000 stakeholders, the Centre has established itself as a trusted repository and collaboration hub. Its long-term sustainability is secured through its integration into the Adra Association, ensuring strategic continuity and expansion beyond the lifetime of the Adra-e project.

In parallel, T3.4 has successfully delivered two virtual ADR Awareness Day events in 2024 and 2025, attracting broad participation across multiple sectors and 16+ countries. These events have provided a venue for expert dialogue on critical topics such as AI trust certification, human-in-the-loop governance, explainable AI, and responsible robotics, while reinforcing the use of WP3 outputs—including educational materials, meta-analytical insights, and trust labelling frameworks. The decision to hold these events virtually reflected a strong commitment to accessibility, environmental sustainability, and EU-wide reach.

Together, these initiatives have:

- Contributed to the strategic goals of Horizon Europe in advancing trustworthy ADR.
- Strengthened the visibility and accessibility of ethical, inclusive, and high-quality ADR education.
- Fostered dialogue and trust across research, industry, and society.

As we transition into the final project phase, WP3 remains committed to scaling these efforts, reinforcing stakeholder trust, and ensuring that the outcomes of Adra-e continue to support Europe's leadership in responsible and human-centric digital transformation.

Appendix – Book 1 and Book 2 Lists of Abstracts

Book 1

Area A – Foundations

- Shaping Europe's Future: The Convergence of AI, Data, and Robotics
  Edward Curry, Fredrik Heintz, Heike Vornhagen, Ahmed Nabil Belbachir, Philip Pietkiewicz, Emanuela Girardi, Marc Schoenauer, Juha Roning.

- The cross-fertilization between the Human-in-the-Loop Approach and the Explainable AI Techniques towards Trustworthiness
  Marina Da Bormida in Cugurra, Sotiris Koussouris, Daniele Crippa, Carl Hans, Robert Hellbach, Andre Tabone, Dimitris Bibikas

  Abstract:
  One of the primary concerns regarding the AI systems regards the potential for them to perpetuate biases and discriminatory practices, since they learn from historical data which may contain inherent prejudices. The human-in-the-loop (HITL) approach, incorporating human expertise and values into AI algorithms and integrating into them the human judgment and oversight, has emerged as a potential solution to tackle with the risks of biased decision-making and related challenges associated with the AI adoption. The AI decisions and actions are supervised and, if necessary, modified or validated by humans: the machine can learn human knowledge and experience during the loop for improving the transparency, the accountability and the performance of AI systems. The HITL approach is helpful for adapting and reconceptualizing the users' role, valorizing their interaction with the learning system, as well as their provision of feedback, guidance, or input when needed. Such a method brings several benefits, making possible the refinement of the AI models, ensuring that they capture the relevant patterns and relationships within the data. Nevertheless, there are also some challenges in the HITL approach, such as the potential slowing down of decision-making processes and the risk of human error. It is therefore important to adopt solutions capable of maximizing the benefits of AI while maintaining a significant and manageable role for human intervention. The Explainable AI (XAI) techniques play a paramount role in this direction, since they make the AI-based decision-making processes and outcomes more transparent and understandable, thereby facilitating a more informed and efficient HITL process, including human oversight. Explainability is therefore key to adopt a human-centered perspective, providing the human operators with the tools needed for the effective supervision and intervention, allowing the users to contribute for improving the AI model. The XAI techniques make the process more accessible and efficient and integrate explanations into it, giving rise to a bi-directional communication channel between the human operators and the intelligent systems, prioritizing the users as the primary driver of this interactive bi-directional process to achieve the desired system behavior, whilst empowering them to collaborate with the systems for enhancing the effectiveness of the interaction and the outcomes of the system, ensuring adaptability. This Chapter explores the HITL approach and the XAI techniques from a double perspective, technical and ethics-driven, outlining how such approach can be effectively combined with automation through the underlying, implicit application of Explainable AI techniques to make transparent all AI models and their associated results. Such techniques also contribute to filtering out (training) data that may hurt the model performance because of poor quality or biases. The synergic use of the HITL approach and XAI can also contribute to prevent or minimize the so-called "hallucination effect", occurring when the AI systems generate inaccurate or logically inconsistent responses, outcomes or data, relying on patterns learned during training that can lead to incorrect or far-fetched conclusions. In these situations, the human intervention is paramount to guide, correct and improve decisions made by machines: the XAI is a key element in navigating the

complexities and challenges like the hallucination effect. The Chapter also outlines the initial findings of the validation of this cross-intersection of the HITL and XAI in the AI-DAPT pilot case focused on robotics and cognitive ergonomics, aiming to optimize working conditions through effective human-machine collaboration. This pilot case will demonstrate how this joint action of the HITL and XAI might be paramount in different scenarios to prioritize the role of the human towards fairness, accountability and trustworthiness.

- Hybrid Intelligence: The fusion of science-based and Machine Learning models
  Eleni Lavasa, Theodora Chrysoula, Vasileios Gkolemis, Charalambos Sinnis, Theodore Dalamagas, Sotiris Koussouris, Nefeli Bountouni, Konstantinos Perakis, Viktor Daropoulos, Charalambos Lambri, George Ioannou, George Pallis, Marios Dikaiakos, Daniele Crippa, Andre Tabone, Dimitrios Bimpikas, Stratos Keranidis, Carl Hans, Robert Hellbach, Dimitris Bouras
  Abstract:
  Process models and parameter estimation have long been fundamental tools in domains such as manufacturing, robotics, power plants, and biomedicine. Early modeling approaches relied on complex systems of mathematical differential and algebraic equations, known as science-based models, to capture process knowledge and scientific expertise. These models leverage physical and chemical properties, static and dynamic behaviors, and causal relationships among observed quantities to support predictive control and operational optimization. As white-box models, they provide transparency by uncovering the inner logic and decision-making steps of the process. However, the advent of Industry 4.0 has brought a surge in available data from industrial processes, driving the rapid growth of machine learning (ML) models. These black-box models excel at discovering patterns and nonlinear relationships in data but often lack interpretability. Hybrid models, which combine the strengths of science-based and ML models, have emerged as a powerful solution. By integrating the transparency and domain knowledge of science-based approaches with the adaptability and predictive capabilities of ML, hybrid models enhance accuracy, robustness, and scalability. This chapter explores the foundations of hybrid models, their development, and applications, providing a comprehensive perspective on their transformative potential across various scientific and engineering domains. This work is done in the context of the AI-DAPT EU project.

- Frugal Machine Learning for Energy-efficient, and Resource-aware Artificial Intelligence
  John Violos, Konstantina-Christina Diamanti, Ioannis Kompatsiaris, and Symeon Papadopoulos
  Abstract:
  Frugal machine learning refers to the practice of developing machine learning models that are efficient, cost-effective, and resource-conscious. This approach aims to achieve high performance while minimizing the computational resources, time, and data required for training and inference. This chapter explores recent advancements, applications, and open challenges in frugal machine learning, emphasizing its relevance in resource-constrained scenarios such as edge computing, IoT devices, and environments with limited bandwidth or high latency. The chapter highlights key technological enablers, including model compression, energy-efficient hardware, and data-efficient learning techniques, as well as their role in fostering sustainability and accessibility in AI. Furthermore, it provides a comprehensive taxonomy of frugal methods, discusses case studies across diverse domains, and identifies future research directions to drive innovation in this evolving field.

- SRIDA: Charting the Future of a Progressive, Inclusive, and Sustainable European ADR Ecosystem
  Fredrik Heintz, Katerina Linden
  Abstract:
  The Strategic Research, Innovation, and Deployment Agenda (SRIDA) is a defining framework for the AI, Data, and Robotics Association, and a forward-looking roadmap for

the advancement of ADR in Europe. It aims to align ADR research and development efforts, address societal challenges, enhance economic competitiveness, and inform the European Commission's Horizon Europe work programme. This chapter outlines the SRIDA's purpose, the stakeholders involved, and the collaborative methodology behind its creation. It describes the operational framework, including tools, contribution processes, development phases, and iterative updates that ensure the SRIDA evolves in alignment with diverse priorities and voices. The SRIDA's continuously updated nature reflects its commitment to transparency, inclusivity, and balancing technical and non-technical goals. Building on previous SRIDA editions, the chapter examines constraints such as balancing agendas, fostering equitable representation within the ADR community, and aligning immediate objectives with long-term strategies. By addressing open challenges and providing actionable recommendations for the future SRIDA iterations, this chapter demonstrates SRIDA's potential as a driver of Europe's ADR leadership, paving the way for sustainable innovation.

- Generative artificial intelligence to tackle visual data accessibility challenges Lore Goetschalckx, Kaili Wang, Siri Willems, Tom De Schepper Abstract:
  Deep neural networks, the cornerstone of the modern AI revolution, had their first breakthroughs on visual tasks (e.g., object recognition). This is not a coincidence given the high interest in such capabilities across industries (e.g., autonomous driving, medical imaging). However, to train, evaluate, and improve vision networks for specific applications, one needs extensive data. Transformative advances in generative AI (GenAI) have recently made it possible to synthesize highly realistic, complex visual data. Harnessing the power of GenAI can overcome challenges faced with real-world data (e.g., scarcity, cost), and ultimately foster more visually capable machines. Here, we offer an overview of the different types of visual GenAI and examine their adoption in real- world applications. We focus on autonomous driving (AD) and healthcare applications because of their high societal impact and unique data challenges (e.g., privacy concerns for patient data, need for synthetic worlds to safely test AD stacks). Finally, we discuss current limitations and promising future directions. This chapter highlights data accessibility solutions as one of many fruits of the industry-transforming technology that is GenAI.

B – Transformations

- The role of Digital Transformation in Manufacturing under the light of AI integration John Angelopoulos, Kosmas Alexopoulos

  Abstract:
  Digital transformation can be realized as a continuous process targeting at the optimization of organizational processes and capabilities, by integrating cutting-edge and AI technologies. In the context of Industry 4.0 and 5.0, manufacturing industry is being revolutionized, yet the implementation/integration of these technologies is followed by significant challenges. Existing practices and solutions do not fully satisfy Industry 4.0 requirements. Further to that, well-established systems are not susceptible to reconfiguration, thus increasing the complexity of the AI focused digital transformation. To add-up to the challenges, recent societal concerns, such as the need for low-carbon transformation to address climate change, require further elaboration under the scope of digital transformation. Extending the above-mentioned concepts to the Industry 5.0, increasing attention is targeted towards aspects such as resilience, particularly during periods of uncertainty or external disturbances, by investigating possible solutions/practices for adaptivity. For traditional manufacturing companies, digital transformation based on the utilization of industrial internet platforms can be realized through three affordances, in particular consonance, resonance, and adaptation. Therefore, by aligning operations with platform capabilities, companies are more capable of implementing more refined practices and advance their digital maturity. From the most pertinent literature it can be extracted that

the technology–organization–environment (TOE) framework will play a key role for reshaping manufacturing by means of complete digitalization, in particular for small and medium-sized enterprises. Despite the evident outcomes of digital transformation on a global scale, tailored strategies are imperative, since “one-size-fits-all” approaches are ineffective. Therefore, this chapter serves as an in-depth analysis of the current industrial landscape under the scope of digital maturity with a focus on AI integration, discussing technological, organizational, and operational aspects. Ultimately, the goal is to identify concurrent challenges, algin them with research and development efforts and structure a roadmap towards digital equality.

- AI-based Management Approaches for Production Processes – Identifying Requirements for Frameworks and Architectures
  Anna Sumereder, Robert Woitsch, Damiano Falcioni, Marlene Mayr Abstract:
  Among the most prevalent challenges for organizations are process transformation, integration and organizational complexity reflecting management obstacles. Processes exist in various formats and formalizations such as textual documentation, graphical instructions or the like. Here, appropriate combinations of AI can be effective to increase transparency and explainability as well as to reduce users’ effort. While symbolic AI is reliable and explainable, sub-symbolic AI is flexible and data-driven. We propose a hybrid AI-based management approach to extract and structure processes in process landscapes. Within individual organizations, these process knowledge repositories enable interactions with AI services and retrieve learnings. Beyond organizations, the suggested management approach entails a wider impact on decision-making by learning from process knowledge repositories across a domain and using it to guide AI results. Showcased by two European research projects: (1) MODPATO reflects on the role of processes in the context of modular production environments. (2) FAIRWork applies hybrid AI-based management approaches to (a) improve information access supporting maintenance and information reliability, (b) validate document compliance, and (c) assist decision-making for worker allocation and production planning. The discussion presents how hybrid AI-based management approaches and results derived from production processes can be further elaborated.

- Data, AI, Robotics transformative power in Industry 5.0. New applications and pilot cases for a more Sustainable and Circular Manufacturing
  Gusmeroli Sergio, Dalle Carbonare Davide, Anibal Reñones, Gema Antequera Garcia, Carmen Polcaro, Cinzia Rubattino, Gabriella Monteleone, Pietro Pittaro, Juanan Arreta, Marielena Marquez Barreiro, Santiago Muños-Landin, Lacalle Ignacio, Jesus Alonso Abstract:
  Industry 5.0 (I5.0) is an evolution of the fourth Industrial Revolution (Industrie 4.0) aiming at a more human-centric, resilient and sustainable Manufacturing. This proposed chapter aims at illustrating the decisive contribution of ADR (AI Data Robotics) technologies in the development, implementation and demonstration of Industry 5.0 principles and in particular of its Sustainable and Circular Manufacturing pillar. In the domain of EC-funded projects (Horizon and Digital Europe programmes HEP and DEP) several initiatives have been recently addressing the I5.0 Sustainability/Circularity challenge from different viewpoints: the technology perspective in HEP ADRA partnership, the application perspective in HEP Made in Europe MiE partnership, the deployment perspective in DEP initiatives such as Data Spaces, AI TEFs (Testing Experimentation Facilities) and EDIH (European Digital Innovation Hubs). The BDVA Smart Manufacturing Industry SMI group is the business-technology community, where the impact of Data Spaces, AI and Digital Twins to Manufacturing Industry is studied, analyzed and promoted. It encompasses more than 130 experts representing more than 60 organizations involved in dozens of EC-funded projects.

- Towards the Irish Mobility Data Space: Challenges, Opportunities, and Requirements for Data Sharing

Rafiqul Haque, Diarmuid Ó Conchubhair, Muhammad Asif Razzak, Majjed Al-Qatf, Fatemeh Ahmadi Zeleti, Wassim Derguech, Edward Curry Abstract: The rapid evolution of digital technologies is transforming the transportation sector in Ireland, with Mobility as a Service (MaaS) emerging as a promising solution to enhance mobility efficiency and sustainability. However, the successful implementation of MaaS in Ireland faces significant barriers, primarily due to the lack of standardized, interoperable, and secure data-sharing frameworks. This research aims to explore the challenges and opportunities of mobility data sharing, and requirements for building a data sharing framework in Ireland. Additionally, it proposes the establishment of the Irish Mobility Data Space (IMDS), a platform for data sharing between mobility stakeholders. Through structured workshops with key private and public stakeholders, this study identifies critical technical, legal and governance-related challenges impeding effective mobility data sharing. By analyzing these insights, we outline the requirements for a robust, sustainable, and user-centric IMDS in Ireland. The proposed IMDS is designed to support critical areas such as MaaS, EV Charging, and Infrastructure planning. It can drive the future of smart and sustainable mobility in Ireland.

- Towards a Holistic Framework for Human-AI Collaboration in Safety-Critical Systems Milad Leyli-abadi, Ricardo J. Bessa, Daniel Boos, Clark Borst, Alberto Castagna, Ricardo Chavarriaga, Duarte Dias, Adrian Egli, Andrina Eisenegger, Joost Ellerbroek, Anna Fedorova, Cristina Felix, Anton Fuxjäger, Joaquim Geraldes, Samira Hamouche, Mohamed Hassouna, Sjoerd Kop, Bruno Lemetayer, Giulia Leto, Roman Liessner, Jonas Lundberg, Antoine Marot, Maroua Meddeb, Manuel Meyer, Hélio Sales, Viola Schiaffonati, Manuel Schneider, Irene Sturm, Julia Usher, Herke Van Hoof, Jan Viebahn, Toni Waefler, Mouadh Yagoubi & Giacomo Zanotti Abstract:
  The interaction between humans and AI in safety-critical systems presents a complex set of challenges that existing frameworks fail to fully address. These challenges arise from the need to balance transparency, trust, and explainability with the imperative for robust and safe decision-making. Currently, no comprehensive framework exists to seamlessly integrate human and AI capabilities while addressing these critical issues, leaving significant gaps in the design, deployment, and maintenance of safe and effective systems. To address these gaps, this chapter will focus on an interdisciplinary approach to develop a conceptual framework tailored for critical infrastructures. This approach bridges traditionally siloed disciplines, including psychology, cognitive engineering, and AI, enabling the study of expert collaborative decision-making in complex, automation-supported scenarios. The framework incorporates design and evaluation criteria that enhance human decision-making, drawing insights from mathematics, decision theory, computer science, and specialized engineering fields such as energy and mobility. Additionally, systems engineering principles and theories adapted for trustworthy AI integration underpin the design of the system's operational, functional, and logical architecture. This contribution aligns with several foundational topics discussed in Part A, including Interaction and Human-Robot Collaboration, Decision-Making, Systems and Methodologies, Interoperability, and Novel Design Concepts.

- CyclOps: Leveraging Semantic Technologies for AI and Data Life Cycle Management and Governance
  Monica Caballero, Anna Queralt, Diego Calvanese, Alexandros Oikonomidis, Chukwuemeka Muonagor, Alberto Abella, Piero Campalani, Benjamin Cogrel, Gabriele Tassi, Julia Palma, Alex Barceló, Saeed Talebzadeh, Ross Little, Peio Oiz, Florencia Pérez, Alexandros Kalafatelis, Maria Font, Sergi Nadal, Alexandros Nizamis, Admela Jukan, Simone Martin Marotta, Ricardo Simón-Carbajo, Javier Conejero, Dolores Ordóñez Abstract:
  Advanced AI data-driven applications have shown the power to reshape our industry.

Despite research advances, the management of the AI and data life cycle in organisations remains largely manual and ad-hoc. In parallel, Data Spaces initiatives are emerging, providing new data sharing opportunities but posing new challenges to leverage all its potential. Tackling these challenges, this chapter describes the new framework proposed by the CyclOps project, to provide interoperable and trustworthy automatic management, governance, and maintenance of the entire data life cycle for large-scale volumes of data generated in heterogeneous distributed sources to enable AI-based data-driven applications for all players, business and research alike. CyclOps operationalizes the end-to-end data life cycle, having in its core knowledge graphs, an established semantic formalism to represent data and metadata adhering to the FAIR Principles, while capturing relevant information to improve reproducibility, traceability and explainability of the AI results. This core layer is complemented with tools for the automation of data management tasks, distributed data processing, AI tools, algorithms, and models, Data Space interoperability, and a human-centric interface. The Chapter presents the key innovative propositions of CyclOps and its underlying technologies as well as use-cases selected to showcase the generalizability of the approach.

- Intelligent underwater perception: current trends and future directions Nicoletta Risi, Gabriel Guimarães Carvalho, Olivier Vasseur, Phil Reiter, Tom De Schepper Abstract:
  Underwater systems have witnessed an increasing demand for enhanced sensing technologies in a wide range of applications, ranging from pipes inspection and ground exploration to hull inspections and rescue operations. Meanwhile, recent technological advances in automation, robotics and AI have set the stage for a new generation of autonomous underwater systems with increased level of autonomy. However, compared to research on, for instance, terrestrial or aerial sensing and autonomous systems, remote underwater environments present additional unique challenges due to factors such as low-lighting conditions, constrained power resources and limited connectivity. In this chapter, we present an overview of the current underwater sensing ecosystem. We review and highlight the complementary nature of the main current sensing technologies, techniques for intelligent information processing, and software frameworks for underwater simulations. We conclude with an outlook on potential promising future directions that could pave the way for a new generation of intelligent and autonomous underwater perception systems. This chapter on intelligent underwater perception, discusses an area with growing interest from both academia and industry.

- Brain-to-Speech: Prosody Feature Engineering and Transformer-Based Reconstruction Mohammed Salah Al-Radhi, Géza Németh, Andon Tchechmedjiev, Binbin Xu Abstract: This chapter presents a novel approach to Brain-to-Speech synthesis from intracranial Electroencephalography (iEEG) data, emphasizing prosody-aware feature engineering and advanced transformer-based models for high-fidelity speech reconstruction. Driven by the increasing interest in decoding speech directly from brain activity, this work integrates neuroscience, artificial intelligence, and signal processing to generate accurate and expressive speech. We introduce a novel pipeline for extracting key prosodic features directly from complex brain iEEG signals, including intonation, pitch, and rhythm. To effectively utilize these crucial features for natural-sounding speech, we employ advanced deep learning models. Furthermore, this chapter introduces a novel transformer encoder architecture specifically designed for Brain-to-Speech tasks. Unlike conventional models, our architecture integrates the extracted prosodic features to significantly enhance speech reconstruction, resulting in generated speech with improved intelligibility and expressiveness. A detailed evaluation demonstrates superior performance over established baseline methods, such as traditional Griffin-Lim and CNN-based reconstruction, across both quantitative and perceptual metrics. By demonstrating these advancements in feature extraction and transformer-based learning, this chapter

contributes to the growing field of AI-driven neuroprosthetics, paving the way for assistive technologies that restore communication for individuals with speech impairments. Finally, we discuss promising future research directions, including the integration of diffusion models and real-time inference systems.

- A Companion Robot Platform for Exploring Technical and Ethical Aspects in Elderly Care
  Lorenzo Boi, Slivia M Massa, Diego Reforgiato Recupero, Daniele Riboni, Rubén Alonso, Michele Cardinali, Elena Ricci, and Alberto Pirni
  Abstract:
  Large Language Models (LLMs) are driving significant advancements across various sectors. Combined with Robotics, they can lay the foundation for a new paradigm in healthcare, particularly in elderly care and companionship. For years, research has focused on the concept of carebots and robot companions. The latest capabilities of LLMs and robotics promise to address many challenges regarding interaction and personalized experience faced by earlier generations of companion robots. In this chapter, we present the Italian TRI-TECH's project ongoing activities and preliminary findings, with a particular focus on the technological, architectural, and design choices that shape its development. TRI-TECH aims to explore the impact of companion robots on the lives of elderly individuals, with a particular focus on assessing the ethical implications through various indicators. The project integrates advanced LLM technologies to optimize interaction and personalize the experience. Various tailored functionalities are implemented, which aim at fully understanding the participants' requests and demonstrating empathy in every stage of the human-robot interaction. The goal is to break down traditional human-robot barriers, creating a more natural and human-like interaction that fosters smooth and reassuring communication, ultimately improving the overall user experience.

- Robotics in Healthcare IoT: and Deep Learning Perspectives
  A'adel Sayyahi, Seyed Enayatallah Alavi, Morteza Jaderyan Abstract:
  The integration of robotics in healthcare is redefining medical services with unprecedented levels of precision, efficiency, and accessibility. Robotics has advanced significantly in areas such as surgical assistance, rehabilitation, and patient care. The emergence of the Internet of Things (IoT) and deep learning technologies has further augmented these systems, offering seamless connectivity, real-time analytics, and intelligent decision-making capabilities. IoT enables robotic devices to communicate with healthcare systems, while deep learning processes vast datasets to optimize robotic performance. These technologies are poised to revolutionize traditional healthcare, addressing challenges like resource limitations, minimally invasive procedures, and chronic disease management, ultimately improving patient outcomes and streamlining clinical workflows.

C – Future Directions

- Roadmaps and Agendas for Research and Innovation in Artificial Intelligence, Data, and Robotics
  Paulo Figueiras, Giorgos Ioannou, Charalambos Lambri, Sangheeta Reji, Elena Mossali, Martina Imarisio Neviani, Sotiris Koussouris, Nefeli Bountouni, Marina Da Bormida in Cugurra, Robert Hellbach, Dimitris Bibikas, Philip O'Brien, Carlos Agostinho, Ricardo Jardim-Gonçalves
  Abstract:
  Artificial Intelligence (AI), Data, and Robotics are transformative technologies that are driving innovation across industries and are addressing critical societal challenges. A comprehensive research and innovation roadmap in this domain is essential to guide advancements, ensure ethical alignment, and maximize societal benefits. This chapter will outline priority areas in these fields, including data processing for AI, synthetic data generation, autonomous systems, human-centric robotics, and explainable, science - guided and adaptive AI, just to name a few topics. It will emphasize the need for robust

frameworks to ensure ethical AI development, addressing concerns around bias, fairness, and accountability. The AI- DAPT roadmap and research agenda highlights cross-disciplinary collaboration, leveraging synergies among AI, big data analytics, and robotics. It advocates for foundational research in adaptive, explainable, science -guided AI while emphasizing applied research for real-world deployment and providing a strategic vision to align public and private sector investments, strengthening Europe's global competitiveness, and promoting trust and inclusivity in AI, data, and robotics technologies.

- Advancing Industrial Collaboration: The Next Generation of Human-Robot Interaction Anibal Reñones, Mireya de Diego, Francisco Melendez, Zoi Arkouli, Nikos Dimitropoulos, Christos Gkrizis, Fernando Castaño, Rodolfo E. Haber and Leire Bastida Abstract: This chapter explores the transformative role of Artificial Intelligence, Data, and Robotics (ADR) in advancing human-centric approaches to manufacturing, aligned with the European Union's Horizon Europe goals of fostering digital competitiveness and supporting the green transition. It highlights the contributions of three EU-funded projects—ARISE, FORTIS, and JARVIS—designed to enhance industrial Human-Robot Interaction (HRI) through innovation and sustainability. ARISE introduces an agile, open-source middleware and a robust SSH framework to ensure scalable, ethical, and sustainable industrial Human-Robot Interaction (HRI). By leveraging cutting-edge platforms like FIWARE and Vulcanexus, ARISE demonstrates the potential for cost- effective HRI deployments across manufacturing, healthcare, and logistics. FORTIS focuses on developing adaptive and human-centric robotic systems, emphasizing safety, reliability, and multimodal interfaces, with pilot applications in assembly and industrial environments. JARVIS advances robot control systems through its integrated Robot Control Module (RCM), enabling multimodal collaboration and enhanced task efficiency in complex manufacturing scenarios such as battery assembly and component manipulation. Together, these projects demonstrate a cohesive vision for advancing collaborative robotics, particularly in supporting SMEs and extending the current boundaries of Human-Robot Interaction (HRI). By extending the state-of-the-art in Human-Robot Interaction (HRI), they will enhance capabilities in tasks that combine human ingenuity with robotic precision, such as service operations, complex industrial processes, and healthcare applications. These initiatives Will validate AI, data, and robotics integration at scale, showcasing innovative solutions for physical and social interactions in diverse work environments.

- Human-AI Interaction and Visualization Perspectives on ADR Kostiantyn Kucher, Magnus Bång, Jonas Lundberg
  Abstract:
  Recent advances in artificial intelligence (AI), data, and robotics (ADR) have pushed the boundaries of the benchmark performance of the respective methods and have already started to change the landscape in various application domains. Some of these domains are mission-critical, with control activity that must match running processes. For those domains, a number of questions and challenges related to safety, robustness, and trustworthiness of AI and ADR methods and models still remain open, especially in the scenarios involving human operators. In this chapter, we provide an overview of human-centered perspectives on ADR with an emphasis on human-AI interaction, interactive visualization, and visual analytics. We explain the relationship of these fields to the related disciplines and fields (including human factors, human-computer interaction, data science, and autonomous systems). We introduce the readers to basic concepts from these fields, discuss how the prior work fits with ADR principles (including examples from prior work on visualization for explainable AI, cognitive systems engineering for joint human-AI control, and evaluation approaches for human-AI decision support systems). We argue that the techniques and frameworks proposed in these human-centered fields can and should be integrated with ADR methods.

Book 2

Fundamentals

- Data Compliance Services: A Key to Unlocking Data Governance Under the AI Act
  Authors: Ivo Emanuilov, Dessislava Petrova-Antonova, Sylvia Ilieva
  Abstract:
  The AI Act requires high-risk AI systems which make use of techniques involving the training of AI models with data to be developed based on training, validation and testing datasets that meet certain (extensive) data quality criteria outlined in Art 10. Organisations seeking to comply with these requirements face a challenging maze of obligations that they should enforce against both upstream providers of (foundation) models or other (non-high-risk) AI systems and downstream deployers of AI systems. Where the general-purpose AI model is available on an open access basis 'as is', providers of high-risk AI systems will have to conduct vendor due diligence to ensure that whatever models their systems rely on meet the data management obligations of Art 10. Similarly, where the general-purpose AI model is available on negotiable commercial terms, the provider of a high-risk AI system might ask the provider of the general-purpose AI system to make specific commercial warranties related to the training, testing and validation data sets used to train the general-purpose AI model. This process can become tedious and expensive over time, and it would require a great deal of investment of time and resources from any provider of high-risk AI systems. To mitigate this potentially negative impact, recital 67 of the AI Act foresees that the requirements related to data governance can be complied with by having recourse to third parties. These third parties offer certified compliance services including verification of data governance, data set integrity, and data training, validation and testing practices. In this contribution, a blueprint for data compliance services is proposed focused on validation of data governance practices, verification of data set integrity, and data training, validation and testing practices, including through data augmentation and data enrichment. A hands-on experience will be provided, advising companies and research organisations on good data governance practices.

- Trustworthy Generative AI in Practice: Automating EU AI Act Compliance Assessment
  Authors: Nataša Krčo, Ioanna Angeliki Kapetanidou, Srdjan Krčo, Alexandros Nizamis, Anja Jakovljević, Konstantinos Votis

  Abstract:

  The European Union has recently established the AI Act, a legal framework that imposes stringent compliance requirements on AI systems. Assessing and ensuring adherence to these regulations manually is often rather cumbersome for AI developers. This chapter addresses this challenge by proposing an approach that automates EU AI Act compliance assessment, significantly contributing to Trustworthy AI. Drawing upon research within the ENACT project, we present an AI Act compliance checker that utilizes Large Language Models (LLMs) to categorize AI models according to the risk levels explicitly defined in the EU AI Act; ranging from minimal to unacceptable risk. Furthermore, it provides LLM-generated compliance insights, including detailed feedback and recommendations on corrective actions. By streamlining these processes, this compliance checker offers a practical solution for AI developers seeking to ensure conformance with AI legislations and ultimately, to build trust in AI and foster responsible innovation. Therefore, this work is fully aligned with the book's topics on Trustworthy Generative AI and regulatory concerns.

- Building Trust in Synthetic Media: Provenance Frameworks for Generative AI
  Authors: Mansura Habiba, Declan Joyce

  Abstract:

  Generative AI is revolutionizing digital content creation, offering unprecedented capabilities in crafting hyper-realistic text, images, and videos. Yet, its rapid proliferation has

significantly undermined public trust in the authenticity of online media. This chapter critically examines emerging frameworks and technological standards such as Adobe's Content Authenticity Initiative (CAI) and the Coalition for Content Provenance and Authenticity (C2PA), which advocate for robust, tamper-proof provenance metadata embedded directly within digital assets. Through detailed case studies, it analyzes how transparency-enhancing methods, including watermarking technologies like Google's SynthID and explicit AI content labelling, can mitigate misinformation risks and restore confidence in digital communication ecosystems. In addition, this chapter explores the importance of AI transparency and governance, the inherent challenges in implementing effective governance frameworks, recent trends, and potential pathways to overcome these challenges. This chapter further evaluates cultural shifts toward transparency, highlighting the growing consumer demand for disclosure of AI-generated media. This chapter brings together cutting-edge technological advancements, strategic policy responses, and exemplary practices from the industry. It presents a dynamic and forward-thinking framework for restoring trust, accompanied by actionable insights designed for researchers, practitioners, and policymakers alike.

- Mitigating Bias and Enhancing Trust in AI-based Systems through a Synthetic Data-driven Framework
  Authors: *Dimitrios I. Fotiadis*
  Abstract:
  The deployment of Generative AI in critical domains necessitates not only technical excellence but also trust, fairness, and alignment with societal values. This chapter builds upon the FAITH (Fostering Artificial Intelligence Trust for Humans towards the optimization of trustworthiness through large-scale pilots in critical domains) project's human-centric approach and our recent advances in synthetic data generation, offering a comprehensive perspective on mitigating bias and privacy risks across the AI lifecycle. We present a structured methodology that integrates bias assessment, synthetic data generation, and fairness-aware performance assessment within a dynamic risk-based governance model. Drawing from real-world use cases in healthcare, we show how demographic disparities can be addressed utilizing either advanced Bayesian modeling techniques or generative-based ones, such as Generative Adversarial Networks (GANs). Based on FAITH's proposed risk-based methodological framework, we propose actionable recommendations for trustworthy design, including the concept of AI model passport and socio-technical threat modelling. Our chapter emphasizes the importance of combining fairness interventions with domain-specific trustworthiness assessments to ensure inclusive, explainable, and ethically aligned AI systems, ready for deployment in real-world settings.

- The Integrity Challenge of AIaaS Model
  Authors: Georgios Spathoulas

  Abstract:

  Generative AI is one of the most significant advancements in recent computer science history. In a couple of years, it has managed to result into specific products and services that are being widely adopted in various domains. The rapid growth of the technology and the rate under which it is being used creates significant concerns with respect to security and privacy of the end users. On top of that, an emerging concern relates to the model under which relevant services are offered to the end user. An AI as a service (AIaaS) model has been defined and adopted mainly because training, deployment and maintenance of large language models is very challenging and while access to such models is restricted, as they are very valuable assets for the companies that have developed them. In AIaaS model, the inference stage of the machine learning workflow happens on the side of the service provider without the user being able to validate that the model he is being offered is the one being used. An important integrity concern emerges, and we need mechanisms that will allow end users to verify the integrity of the service provider offerings and at the

same time will force service providers to offer to end users the services promised to them. The present chapter present approaches that can support users to that end and discusses relevant challenges and the limitations.

- Generating Bias? Auditing Fairness in Synthetic Tabular Data from Generative AI
  Authors: Swati Swati, Arjun Roy, Emmanouil Panagiotou, Eirini Ntoutsi

  Abstract:

  As Generative AI expands beyond text and image generation into structured data applications, synthetic tabular data is increasingly used for benchmarking, simulation, and privacy-preserving analytics. However, the fairness implications of such data are often overlooked. This chapter investigates whether generative models produce tabular data that preserves, mitigates, or amplifies bias across sensitive attributes such as gender, age, or race. Using lightweight prompting methods with open-source language models, we generate synthetic versions of widely used fairness datasets and apply a multi-dimensional auditing framework to compare real and generated distributions. The analysis reveals how prompt design and model behavior can influence group-level disparities in downstream use. We canvas high-stakes domains like finance and healthcare to demonstrate the societal relevance of these disparities. By highlighting hidden risks and offering empirical insights, this chapter contributes to the responsible use of generative AI in structured data settings. It directly supports the edited volume's goal of advancing reliable, transparent, and human-centric generative AI systems.

- Opening the Black Box: Explaining Generative AI Processes with Scene2Model
  Authors: Wilfrid Utz, Iulia Vaidian, et al.

  Abstract:

  For Generative AI to be trustworthy by domain experts, the processing capabilities (specifically related to reasoning and explainability) need to become transparent. At this stage, approaches like neurosymbolic techniques are applied to combine the strength of both worlds – data-driven processing and knowledge representation, however the domain expert needs to trust that the system is setup and trained according to pre-defined requirements. Therefore it is not fully understood how Generative AI reaches its output, particularly for users not familiar with the underlying technology. To remedy this, we propose a hybrid co-creation approach, namely the Scene2Model platform (https://scene2model.omilab.org) realized by OMiLAB in the context of the FAIRWork project (https://www.fairwork-project.eu), to be utilized. By using the Scene2Model tool, the highly complex technical processes employed by Generative AI can be visually defined using interactive conceptual models, thus enabling users to see how it is intended that a specific output can be reached. This way, the “decision-making” process of a hybrid AI system is explained to the users without first requiring them to gain a deeper understanding of the technology itself. As the use of AI becomes more common across a wide range of domains, there is an increased need for explainability, to heighten transparency and prevent users from trusting outputs reached by flawed procedures. The Scene2Model platform and its tailored approach introduced in this chapter allow AI users to effectively “open the black box” of Generative AI decisionmaking, thus increasing trustworthiness by removing the barriers to understanding the internal technical processes of the AI system. In this chapter we explore and demonstrate in a prototypical environment the design process, artefacts and outcomes to support intelligent service interactions using the Scene2Model platform. Evidence and evaluation results from the project are presented.

- Responsible Development of Generative AI for Societal and Economic Resilience
  Authors: Michael McCann, Muslim Jameel Syed

  Abstract:

Generative artificial intelligence (GAI) is transforming our ways of business, communication, learning, and decision-making. Alongside its remarkable potential, this technology presents significant obstacles that relate to technological, ethical, and societal issues, including trust, fairness, bias, and its effects on individuals and society. As GAI becomes increasingly embedded in our lives, ensuring its reliability and that it serves the public good has become crucial. This chapter explores what are the current design technical challenges and how ethical principles and regulatory frameworks can work together to guide its responsible development, deployment, improving economic growth, protecting human rights, promoting inclusion and diversity, and building systems we can trust and use for nation prosperity. This chapter examines core design and ethical principles, including transparency, accountability, inclusion, and human oversight, and evaluates their implementation across the AI lifespan. Simultaneously, we examine the developing regulatory frameworks, such as the EU AI Act and GDPR, to comprehend their function in assuring compliance and safeguarding public interest. Emphasis is placed on economic risks, including design support for small and medium enterprises, cross-border trade, and legislative compliance, as well as societal risks such as AI-enabled cyberbullying, misinformation, and psychological harm, particularly affecting women, children, and marginalised communities. This chapter offers pragmatic perspectives for developers, politicians, and organisations aiming to harmonise innovation with ethical accountability and legal adherence in the era of Generative AI.

- AI Bill of Materials (AI BOM) for Trustworthy Generative AI
  Authors: Mayuko Kaneko

  Abstract:

  The proliferation of Generative AI models has raised significant concerns regarding their potential for hallucination, bias amplification, copyright infringement, regulatory non-compliance, and other risks. This chapter investigates the application of an AI Bill of Materials (AI BOM) framework as a mechanism to mitigate these challenges of Generative AI. Building upon existing AI BOM research, including model cards and data cards, we propose a refined and specialised AI BOM approach for Generative AI. Specifically, we argue that a comprehensive AI BOM for Generative AI must incorporate elements such as specifications for prompt engineering, provenance tracking of generation history, and rigorous, quantifiable output evaluation methodologies. The AI BOM for Generative AI should also address the critical issue of reproducibility through meticulous management of dynamic elements and version control. This chapter outlines key requirements inherent in developing such a framework. Furthermore, we present preliminary findings and practical challenges encountered during an ongoing experiment focused on the creation of an AI BOM for Generative AI, offering valuable insights and highlighting open research questions for developers and researchers seeking to enhance the trustworthiness of Generative AI.

- Cognitive Load, Safety, and Generative Co-Pilots: Insights from Engineering and Neuropsychology
  Authors: Mario Espinosa Gámez

  Abstract:

  Generative AI systems are being introduced as cognitive co-pilots in safety-critical applications across transportation, medicine, and energy control systems. Their design and integration raise fundamental questions of mental workload, human performance, and system safety. This chapter discusses the intersection of cognitive ergonomics, engineering psychology, and AI interaction design to review the promises and challenges of these systems. The chapter begins with a description of cognitive load theory and how it applies to high-stakes, high-information contexts. It then connects these results to neuropsychological findings regarding attention, working memory, and multitasking. Specific attention is placed on how generative AI can both diminish and unintentionally

heighten cognitive pressure. A key section of the chapter addresses automation surprise, overtrust, and impaired vigilance risks. Drawing from empirical evidence in aviation and medical systems, it explains design flaws that cause operator errors. On this basis, the chapter presents design principles for adaptive, transparent, and cognitively aligned co-pilots. The chapter concludes by providing recommendations on system testing, interface adaptation, and policy frameworks that ensure human-centered safety in the implementation of generative AI. The proposed contribution aligns with the volume's focus on human-machine interaction and offers practical and theoretical contributions on neurocognitive safety in emerging AI systems.

- Generative AI and the Erosion of Cognitive Autonomy: A Neuroethical Framework
  Authors: Mario Espinosa Gámez

  Abstract:

  Large language models and other generative AI models increasingly perform jobs that were once the sole domain of human cognition – composing, summarizing, solving, and deciding. While these technologies deliver marked improvements in productivity, they also raise novel ethical concerns regarding the risk of eroding cognitive autonomy: the ability of individuals to reason for themselves, think for themselves, and decide for themselves. This chapter will provide integrative neuroethical analysis of cognitive autonomy during the age of generative AI. Based on neuroscience, cognitive psychology, and ethical philosophy, it will debate how over-dependence on AI can reconfigure users' mental engagement, memory functions, and epistemic responsibility. It will also take into account ethical risks like automation bias, epistemic passivity, and vulnerability of particular user groups. Building on this foundation, the chapter will subsequently propose a normative model for the design and regulation of autonomy-maintaining systems. It will conclude with practice policy and design recommendations aimed at finding balance between individual agency and cognitive support. The contribution is in the scope of the edited book as it offers a multidisciplinary and futurophilic perspective in AI ethics and neurotechnology in understanding how cognitive integrity can be preserved in more automated human-AI interactions.

Applications

- Responsible AI in Practice: Designing a Trustworthy GenAI Chatbot in Public Administration
  Authors: Gizem Gezici, Chandana Sree Mala, Fosca Giannotti, Sezer Kutluk

  Abstract:

  This chapter presents a case study with best practices for the development of a trustworthy generative AI chatbot within the public administration domain, specifically designed to support the employees of Regione Toscana—the regional government of Tuscany, Italy, responsible for delivering public services and managing digital infrastructure across departments. The system aims to improve the accessibility and usability of internal ICT services of Regione Toscana by translating complex technical content into natural language responses tailored to non-specialist users. Positioned within the “limited risk” category as defined by the EU AI Act’s risk-based framework, this proof-of-concept addresses critical reliability concerns associated with large language models, particularly hallucinations—plausible yet inaccurate or ungrounded outputs. To mitigate these risks, the system mainly integrates two key safeguards: (1) source transparency, allowing users to trace responses directly back to the original documentation, and (2) authoritative grounding, ensuring that all outputs are derived from verified and up-to-date institutional sources. The chapter discusses the architectural choices, alignment with ethical and regulatory requirements, and design considerations for human-centric AI in public sector contexts. It contributes to

ongoing discourse on responsible generative AI deployment by demonstrating how technical safeguards can support transparency, compliance, and user trust in environments involving non-expert users

- Trustworthy Generative AI for Unmanned Vehicle Swarms
  Authors: Mai Le, Barry O'Sullivan, Hoang D. Nguyen

  Abstract:

  The integration of artificial intelligence (AI) into multi-unmanned aerial vehicles (a.k.a., UAV swarms) has garnered significant attention in both academia and industry due to their potential proving the services that are difficult and dangerous for humans. Generative AI (GAI) has emerged as a powerful AI paradigms, thanks to its powerful language understanding and generalization of diverse digital content, offering great potential in addressing complex problems in UAVs systems. Most GAI techniques have primarily focused on GAI-based applications and technical Problems, rather than examining the accountability and trustworthiness of GAI models, a crucial evaluation criteria for AI human beings. This chapter aims to provide a scoping review of the trustworthiness of GAI in multi-agent UAV systems. Firstly, we present the fundamentals of GAI and representative GAI models, the concept of multiagent systems and evaluation methods trustworthiness of GAI systems. Secondly, we review innovative techniques that address trustworthy requirements in terms of safety, robustness, privacy, accountability, explainability, and fairness, along with challenges in multi-agent UAV networks. Finally, we highlight several promising solutions and future research directions.

- From Bench to Bedside: Auditing Trustworthiness of Generative AI in Healthcare Practice
  Authors: Alin Navas

  Abstract:

  This chapter addresses the challenge of developing and deploying trustworthy generative AI—especially large language models (LLMs)—in healthcare. We begin by outlining robust methods to identify and mitigate bias, including data-level interventions (such as diverse data curation and augmentation), model-level techniques (like fairness constraints and fine-tuning), and post-processing adjustments (such as output debiasing through prompt engineering). The chapter then explores explainability in LLMs mainly through model introspection (e.g., attention analysis, probing), and human-in-the-loop paradigms that allow clinicians to interrogate model outputs and reasoning. We examine the current landscape of LLM evaluation, highlighting the limitations of traditional benchmarks and the emergence of comprehensive, clinically relevant evaluation frameworks (such as MedS-Bench, MedHELM, HealthBench) that measure performance across real-world medical tasks and multiple metrics, including accuracy, faithfulness, and robustness. We then emphasize the need for AI literacy among physicians. This includes foundational understanding of AI principles, the ability to critically assess and validate AI outputs, and readiness to collaborate with data scientists. Finally, we survey common healthcare applications of generative AI, from ambient AI scribes and diagnostic support to medical imaging report generation, administrative automation, and medical education.

- Trustworthy Generative AI for Real-Time Travel Risk Management: Personalised Safety at Scale
  Authors: Dr Sheelagh Brady

  Abstract:

  This chapter presents **Kowroo**, a real-time travel risk platform that applies trustworthy AI principles to enhance individual safety and organisational duty of care. Combining generative and predictive models with contextual, crowdsourced, and deeply personalised data—including user profiles, location, travel intent, and perceived safety indicators—Kowroo delivers tailored risk assessments and proactive alerts in dynamic environments.

Within the chapter, we explore how such personalisation offers both opportunity and responsibility: improving relevance and responsiveness while raising ethical challenges related to bias, overreach, and privacy. The chapter critically examines how the system balances transparency, fairness, and privacy, with a focus on safeguards such as explainability, opt-in controls, and data minimisation while delivering timely, actionable insights through human-centric design. Drawing on live trials, we contribute a real-world case study from a European AI startup navigating the implementation of generative AI in a safety-critical, perception-sensitive domain. We offer a pragmatic framework for AI practitioners, regulators, and designers seeking to build systems that responsibly manage sensitive personal data, foster user trust, and uphold human-centric values.

- Advancing User-Centric Manufacturing in High-Impact Industrial Sectors through Multimodal GenAI
  Authors: Sotiris Makris

  Abstract:

  Despite significant research progress, the wide industrial adoption of Human-Robot Collaboration remains hindered by challenges such as high robot programming complexity, low performance of human-robot collaborative tasks, and limited operator acceptance. Recent advancements in Generative AI have the potential to alleviate such barriers in several ways. This chapter introduces the JARVIS GenAI framework which relies on state-of-the-art Large Multimodal Models and Probabilistic/Neural Generative Models to enable accurate robot perception, natural, context-aware human-robot interaction, dynamic robot task adaptation, and intuitive robot programming. The proposed framework is tested in two high-impact industrial sectors of agile manufacturing, specifically: a) assembly of hybrid cars battery packs (automotive), b) passenger aircraft seats manufacturing (aeronautics). In both settings, human-robot collaboration is central, involving multiple product variants, tools, and resources—making real-time task allocation, process control, and seamless interaction particularly demanding. The JARVIS GenAI framework shows promise for yielding higher efficiency in onsite operator support, quicker and easier robot programming and robot operations configuration, more intuitive interaction, improved operator acceptance, and more reliable robot perception.

- Bridging the Trust Gap: Hybrid Decision Support Systems as a Pathway to Trustworthy Generative AI
  Authors: Artur Bogucki

  Abstract:

  Artificial Intelligence (AI) has transformative potential to enhance human decisionmaking, mitigate cognitive overload, and reduce bias, particularly in high-stakes scenarios. Yet, the adoption of AI-based decision support remains minimal due to challenges in assessing AI assumptions, limitations, and intentions. This creates a critical gap in realizing trustworthy generative AI systems that individuals can reliably depend on, assured that AI 'understands' their needs and genuinely aims to assist them. This chapter examines how TANGO project's hybrid AI approach addresses this gap through a custom ethical, legal, and socio-economic (ELSE) assessment methodology. This method systematically evaluates hybrid decision support systems (HDSS) across four diverse, high-impact case studies: prenatal and postpartum care, surgical decision-making, credit lending, and public policymaking. Key findings indicate that integrating cognitive theories of mutual understanding, cognition-aware explainability, and a human-in-the-loop model significantly enhances AI reliability, transparency, and alignment with human values. Case studies evaluations highlight crucial lessons: effective human-AI collaboration requires tailored information delivery, acknowledgment of intuitive versus deliberative decision-making, and ongoing, interactive learning. Implications suggest that the TANGO framework provides fundamental guidance for developing and applying generative AI systems rooted firmly in trustworthiness.

Consequently, it positions Europe as a global leader in human-centric AI development, setting benchmarks for future innovations.

- Generative Models in Medical Imaging: when Less is More
  Authors: Khaoula ElBedoui, Walid Barhoumi

  Abstract:

  The rise of Generative AI (GAI) in medical imaging marks a pivotal shift in the healthcare imaging pipeline by addressing input problems, enhancing model performance, and improving output interpretation. By correcting data limitations, increasing diagnostic accuracy, and streamlining report generation, generative models emerge as critical enablers of trustworthy and clinically meaningful AI. In fact, at the data level, GAI models break traditional constraints by synthesizing realistic data, enhancing image quality, generating missing modalities, and balancing class distributions. At the model level, GAI improves medical tasks such as lesion segmentation, disease classification, and anomaly detection, based on the use of advanced architectures and new learning paradigms. Finally, manual reporting bottlenecks and variability in clinical summaries can be overcome by automating finding generation and improving results interpretability and explainability. Grounded in the principle that “Create More, Consume Less” or “Less is More”, we highlight the contribution of GAI in critical imaging tasks. In this study, we outline the contribution of generation design, under ethical constraints, full transparency, and clinical validation. By doing so, we argue that GAI paves the way for 5P Medicine (Preventive, Predictive, Personalized, Precision, and Participatory) enabling new frontiers in real-world healthcare settings with effective and actionable solutions.

- Exploring Diagnostic Innovations: Survey of Deep Learning Models for Lung Disease Classification
  Authors: Hoda El-Batrawy, Wesam Ahmed

  Abstract:

  This study meticulously investigates the diagnostic efficacy of infection categorization within a spectrum of advanced computational models, emphasizing the enhancement of initial diagnostic precision. Deep learning, heralded as a transformative paradigm in automated detection, has profoundly reshaped diagnostic medicine, particularly in addressing infectious diseases. This comprehensive review consolidates a meticulously curated corpus of literature, focusing on the application of medical image classification techniques—primarily chest radiographs—for the detection of infections such as pneumonia and COVID-19. It delves deeply into the methodological architectures employed across studies, encompassing cutting-edge detection frameworks, sophisticated classification algorithms, datasets of varying complexity, and their resultant performance metrics. These findings are rigorously evaluated, systematically compared, and analytically articulated to elucidate their contributions to the advancing domain of medical diagnostics. Furthermore, the review presents a critical comparative analysis of extant survey papers, highlighting their methodological depth, inherent constraints, and potential for misuse. By discerning the strengths and limitations of these methodologies, this work provides invaluable insights and delineates strategic pathways for the progressive evolution of intelligent diagnostic systems and the augmentation of clinical decision-making paradigms.

- Manufacturing, Automotive, Engineering, and Industrial Sectors, including Generative AI for Robotics
  Authors: Mahmood Seyyedzadeh

  Abstract:

  This chapter provides a comprehensive survey of generative artificial intelligence’s transformative role across the manufacturing, automotive, engineering, and industrial sectors, with a special emphasis on robotics applications. We begin by framing the chapter

within the context of Industry 4.0, defining essential terminology and outlining how generative models—ranging from diffusion processes to transformer-based architectures-are trained, integrated, and extended with hybrid neuro-symbolic techniques for CAD/CAE pipelines. Next, we characterize the digital drivers reshaping traditional industries and detail the data requirements, preprocessing strategies, and system-level integration needed to deploy these models effectively. In the manufacturing domain, we explore generative design for part and component optimization, predictive maintenance through synthetic failure data, process scheduling, and synthetic-data-augmented quality control. The automotive section examines AI-powered digital twins for autonomous-vehicle simulation, generative lightweight-structure exploration, and natural-language interfaces for diagnostics and personalized in-vehicle experiences. Moving to broader engineering and industrial automation, we discuss AI-driven system architecture synthesis, collaborative human–machine workflows, virtual commissioning and simulation, and methods for ensuring safety and reliability. The chapter then delves into generative-AI-driven robotics, covering learning from demonstration, motion planning with generative priors, synthetic sensor-data generation for perception training, and real-time adaptive control. Throughout, we address trustworthiness, focusing on transparency, explainability, data provenance, security, and ethical considerations, and introduce evaluation metrics and benchmarking methodologies to quantify performance, robustness, and business ROI. Finally, we identify key challenges and future directions, such as scaling to high-volume production, cross-domain generalization, edge-AI integration, and workforce upskilling, and offer a roadmap for researchers and practitioners aiming to harness generative AI to build the next generation of intelligent industrial systems.

- Generative AI for Robotics
  Authors: Mahmood Seyyedzadeh, Amir A. Ghavifekr

  Abstract:

  In the development of intelligent autonomous systems, the incorporation of generative artificial intelligence (AI) into robotics represents a paradigm shift. This chapter explores the application of generative AI models such as generative adversarial networks (GANs), variational autoencoders (VAEs), and large language models (LLMs) in shaping robotic behaviors, including perception, reasoning, planning, and their interaction with the environment. We review recent developments in data generation for simulation and training, policy generation for control and decision-making, and multimodal understanding for human-robot interaction. It focuses on how transformer-based architectures and foundation models help robots learn with little supervision, generalize across tasks, and adapt to modified situations. Challenges such as safety, interpretability, real-time implementation, and ethical considerations are also covered. The chapter concludes with a perspective on the potential of generative AI to empower the next generation of cognitive and collaborative robotic systems.