

# Case Statement

## EOSC-Future & RDA

### Artificial Intelligence & Data Visitation

### Working Group

26 August 2022, Version 5.0

#### Table of Contents

Table of Contents .....	1
1 Overview .....	1
2 Charter .....	2
Vision .....	3
Mission .....	3
3 Description of Work .....	3
4 Objectives.....	3
5 Deliverables.....	4
6 Methods .....	5
7 Organization and Activities .....	5
8 Background & Founding Members .....	6
9 Membership.....	8
10 Engagement with EOSC-Future, the RDA Community and Other.....	8
11 Timeline.....	8

#### 1 OVERVIEW

The EOSC-Future & RDA Artificial Intelligence and Data Visitation Working Group (AIDV-WG) will address ethical, legal, and social challenges of Artificial Intelligence (AI) and Data Visitation (DV) affecting of state-of-the art data technology impacting scientific exchange in the context of data sharing and the European Open Science Cloud (EOSC). The AIDV-WG has been established through a competitive [call for proposals](#) for RDA Working Groups focusing on the development of solutions for the European Open Science Cloud ([EOSC](#)), working in conjunction with the European Commission-funded project [EOSC Future](#).

## 2 CHARTER

AI and DV are relatively new technologies that have the potential to dramatically facilitate researchers' work across wide and diverse datasets, both within disciplines and across disciplines. Until recently, the focus has been on data sharing (DS), which generally requires a preliminary movement of data to a central repository or a central system of repositories prior to being able to perform an analysis across the disparate datasets. DV introduces the ability to query and perform data analysis across disparate datasets, perhaps even with differing operating, management, and stewardship systems among the various datasets. The conventional DS approach is time consuming and expensive, usually requiring a considerable amount of data reformatting and editing. DS also creates large databases that impose centralized formatting and management, while also potentially containing large amounts of duplicate or redundant data. Moreover, there are issues regarding data synchronization, data management, data sustainability, and data ownership as well as administrative or legal challenges when data crosses borders, be those institutional or juridical borders.

The movement of data creates a multitude of challenges without always solving fundamental problems, such as those of access or interoperability. DV is an alternative approach that aims to alleviate many of these issues. Open science platforms, such as EOSC, have already facilitated researchers' access to one another's work and data. DV is now seen as a step forward toward further empowering the reach and efficacy of Open Science by facilitating a researcher's ability to navigate and access far greater amounts of data, data that is perhaps fragmented across various clouds, registries, networks, and operating systems. Via DV, AI algorithms developed to respond specific research questions travel to the data for analysis instead of first requiring the data to be transported to, and made to fit, a central repository.

Both AI and DV open a myriad of ethical and legal issues, many of which are already present in conventional data sharing. These ethical and legal issues, however, take on new meaning when new technologies and new modes of interaction and conduct are developed for data exchange. These challenges include those of access, privacy, confidentiality, consent, accountability, transparency, accuracy, integrity, and explainability. The AIDV-WG will focus on the ethical and legal issues of these technologies in the context of Open Science. It will also examine the impact of these technologies on society and science policy. As European and global scientific governance moves at an accelerated pace to reform science through Open Science, it is critical that we appreciate the complexities of the ethical, legal, technical, and societal issues within the framework of Open Science and its goals for the future of science in our communities and globally.

The AIDV-WG will build an expert cross-disciplinary network of that includes expertise in AI, data science, and data stewardship as well as involving researchers focused on the ethical, legal, and social questions and challenges that surround these technologies. There is a special interest to include AI designers and developers so that technical aspects of these technologies represented in the work will be strengthened. To embrace diverse perspectives, experts are drawn from across geographical regions. The AIDV-WG will also include experts from ongoing EU and international projects in AI and DV, as well as experts in building technical platforms and enterprises employing either AI or DV, particularly those associated with EOSC or other Open Science platforms. We will consider the need for an EOSC Exchange that enables data transfer, research data as a service, and understands the current research software that would enable or contribute to such an EOSC Exchange. The focus is on building a multidisciplinary understanding of the ethical, legal, and social issues in relation to systems

technologies and system users. We will identify the problems and build our ethical, legal, and social guidance on responses that support the future of the EOSC.

#### **VISION**

The vision of the AIDV-WG is to bring together expertise across disciplines and regions to ensure ameliorate the use of AI and DV in research and innovation across technologies and sectors to address the grand challenges of society.

#### **MISSION**

The mission of the AIDV-WG is to contribute to building the ethical, legal, social, and technical frameworks and bridges enabling the open sharing and re-use of data in the framework of Open Science.

### **3 DESCRIPTION OF WORK**

The AIDV-WG will facilitate the development of tools, policy, and practices that promote EOSC Services (PIDs for EOSC, EOSC Marketplace, PID policies and compliance, AAI), ESOC Exchange (data transfer, research data as a service), and the EOSC Interoperability Framework (EOSC Interoperability Framework, ways of implementing the interoperability framework).

The AIDV-WG will also consider the need for training and capacity-building in how the guidances can be used and implemented.

The AIDV-WG will promote and include the following principles:

The [FAIR](#) Data Principles: Findable, Accessible, Interoperable, and Reusable

The [TRUST](#) Principles for Data Repositories: Transparency, Responsibility, User focus, Sustainability and Technology

The [CARE](#) Principles for Indigenous Data Governance: Collective Benefit, Authority to Control, Responsibility, and Ethics.

The [RDA Guiding Principles](#): Openness, Consensus, Inclusive, Harmonization, Community-driven, Non-profit and technology-neutral.

The [RDA Code of Conduct](#)

The AIDV-WG will strive to facilitate data interoperability nationally and internationally in ways that support the [Sustainable Development Goals \(SDGs\)](#) across their 17 global goals and associated 169 targets.

Working to support the EOSC Future project and facilitate the implementation of EOSC across research communities, this AIDV-WG will examine interoperability issues arising across federated and non-federated systems. Particular attention will be given to national and institutional policies (ethics/legal) and how they affect the generation of metadata and interdisciplinary work and cooperation. Of importance will be to examine the ethical, legal, and social issues arising in EOSC Science clusters with their inter- and cross-disciplinary questions.

### **4 OBJECTIVES**

The principal objective of AIDV-WG is to examine the promises, challenges, and barriers to the use of AI in data sharing and Open Science having regard to scientists and research

institutions as well as to policy and the interests of patients, communities, health advocates, and those stakeholders otherwise underrepresented in these important initiatives for Open Science. The intention is to further develop the objectives of EOSC Future and RDA within the EU and global contexts in these important and related areas with ethicists, legal experts, scientific and research institutions, as well as with patients, communities, health advocates, and those instances otherwise underrepresented in these important initiatives for Open Science.

The following specific objectives will be pursued by the AIDV-WG:

1. Developing an understanding of the ethical and legal challenges AI and DV in Open Science across user groups.
2. Contribute to clarity in, and the development of, the ethics and legal frameworks to facilitate user engagement in Open Science.
3. Provide a European and global platform for engaging ethics, law, and social considerations in AI, data sharing, DV, and Open Science policy development.

## 5 DELIVERABLES

The project's deliverables will be developed to support broad understanding of the underlying values of the EOSC and RDA. These deliverables will be designed for use to develop capacity, particularly regarding competence building across skill sets while also contributing to training programmes, in the EU and globally. Through its inclusive and open design, the AIDV-WG's outputs to EOSC as well as other national and international open research commons initiatives. Our working methods are based on co-creation and cultivation, founded on the following principles: Trust - Shared Vision - Leadership - Open Communication - Democratic Engagement - Clear Roles - Goal Driven - Growth/Vibrancy - Standards and Processes - Discovery Enabling - Resourcefulness.

The AIDV-WG will pursue the following deliverables:

- 1. A survey on current ethical, legal, policy, and societal frameworks for AI and DV:** The survey will explore these considerations from within EU institutions as well as in other national and international institutions to assess the key challenges and gaps in AI and DV, having regard to the needs of researchers and societies in Europe and globally taking into account SDG-associated organizations.
- 2. Guidance on legal considerations for AI and DV:** a mapping of legal considerations for AI and DV as well as how to navigate legal frameworks for users of EOSC and other Open Science platforms. Specific attention will be given to the Schrems II decision and the effect of this on DV.
- 3. Guidance for informed consent in AI and DV:** The GDPR and other EU data and AI regulations as well as regulations in other jurisdictions have placed heavy emphasis on the role of informed consent in data sharing and data publication. We will examine the role of informed consent in AI and DV, addressing fundamental challenges to current informed consent frameworks and practices. The aim is to provide guidance for researchers and data controllers across disciplines regarding informed consent in AI and DV.
- 4. Guidance for ethics committees reviewing AI and DV:** Ethics committees (RECs/IRBs/IECs) have been confronted by new challenges when encountering the need for advice on data management and data sharing as well as in other areas of data processing. The use of AI and DV, especially in health-related research, requires investigation with regard to the ethical, legal, and social issues these raise for ethics committees and those submitting

proposals for advice/approval to ethics committees. This guidance will assist ethics committees in understanding questions, methods, and procedures for reviewing AI and DV.

**5. An AI Bill of Rights:** Underlying the growing application and use of AI and DV is a concern to ensure that data subjects are protected by these new technologies. The AIDV-WG will draft an RDA AI Bill of Rights that promotes fundamental human rights and advances trust in AI and federated systems for Open Science.

## 6 METHODS

The AIDV-WG will work within the EOSC-Future and RDA frameworks promoting

- Organizational collaboration
- Data sharing and interoperability
- Data sharing and data governance
- Capacity building

The following methods will be used by the AIDV-WG:

1. Monthly Zoom meetings (first Monday of each month for 18 months)
2. Focused meetings
3. Running agenda items, including reports of ongoing projects and additional items added given upcoming events or emerging opportunities aligned with deliverables or collaboration with partnering organizations
4. Co-chairs may also meet outside of these bi-monthly meetings to manage ongoing deadlines
5. Plenary sessions may be organised to provide public updates on the working group's progress
6. activities and space for community feedback regarding strategy or scope.
7. Mentoring: one or two junior researchers will be assigned to each workstream to be mentored by a leader of the workstream
8. Workstreams per deliverable
9. Webinar and meeting participation: to learn and share
10. Publication of each deliverable using open access platforms/journals
11. Sharing of survey materials and data according to the FAIR Principles and Open Science
12. Dissemination of the deliverables through EOSC Future, RDA, and other international data and AI groups as well as through (social) media

## 7 ORGANIZATION AND ACTIVITIES

The AIDV-WG will meet monthly over on the first Monday of each month from 14:00 to 15:00 CET beginning on Monday, 30 May 2022 (exceptionally due to scheduling conflicts) through Monday, 8 January 2024. The co-chairs will share responsibility for developing the meeting agendas, organizing meetings, leading discussions, keeping track of the work and deadlines, and arranging participation in and plenary sessions. Specific task groups will be established to progress on the identified deliverables with regular reporting. A mailing list will be used for asynchronous communications. Other tools may be used for collaborating on and tracking deliverables.

These meetings will frame the following activities:

**Case Statement**  
**EOSC Future & RDA**  
**Artificial Intelligence & Data Visitation Working Group**

1. Surveying community and users' interests and concerns with AI and DV in Open Science frameworks.
2. Workstream sub-meetings to address specific workstreams
3. A mentorship program
4. An outreach program
5. Webinars and conferences by provided by AIDV-WG or participation in such by its members
6. Conduct outreach to outside organizations and experts to gain support on outputs
7. Communicate to EOSC Future and the RDA Groups and Secretariat the importance of adopting the deliverables, together with AIDV-WG members, actively look for adopters.
8. Provide the ERA Secretariat with a publishable version of the AIDV-WG's deliverables, a maintenance plan, contact details of two early adopters, and other information required for publishing the deliverables and for the endorsement process.
9. The work of the AIDV-WG will include the following acknowledgement:  
'This work was developed by the EOSC-Future & RDA Artificial Intelligence and Data Visitation Working Group (AIDV-WG) as part of the European Commission funded European Open Science Research – Future Project and the Research Data Alliance (RDA). We acknowledge the support provided by the EOSC-Future and RDA communities and structures.'

## **8 BACKGROUND & FOUNDING MEMBERS**

The EOSC Future / RDA AIDV-WG develops out of the earlier RDA COVID-19 Legal/Ethical Working Group. This working group examined the legal and ethical frameworks and challenges for the identification, curation, and sharing of (personal) data during the COVID-19 pandemic. It examined policy and best ethics and legal practices while also providing guidance for researchers on data sharing / data exploitation. It made valuable contributions to the [RDA COVID-19 Recommendations and Guidelines on Data Sharing](#) and published an important paper on the GDPR: Hallinan, D. et al. '[International transfers of personal data for health research following Schrems II: a problem in need of a solution](#)'.

The following members of that working group now continue with the AIDV-WG:

**Francis P. Crawley** (Chairperson) is a philosopher specialized in research ethics and methodology, integrity, and data/AI ethics & law. He is the Executive Director of the Good Clinical Practice Alliance – Europe (GCPA) and the Strategic Initiative for Developing Capacity in Ethical Review (SIDCER) located in Leuven, Belgium. He has been a member of the Research Data Alliance's (RDA) Legal and Ethics Working Group and the Community Participation Working Group. He recently launched and chairs the Artificial Intelligence and Data Visitation Working Group (RDA AIDV-WG). He is also a member of the Research Data Publishing Ethics Working Group with FORCE11 & The Committee on Publication Ethics (COPE), the Pistoia Alliance, Regulatory and Ethics Work Stream of the Global Alliance for Genomics & Health (GA4GH-REWS), and the Data Stewardship Working Group of the Virus Outbreak Data Network (VODAN) GO-FAIR and a member of the Ethics Working Group of the International Federation of Associations of Pharmaceutical Physicians (IFAPP). He brings 30 years of experience working closely with patients, communities, researchers, and policy-makers across disciplines, domains, and geographic regions in establishing consortia, developing patient registries, contributing to the development of biobanks, drafting data management and data protection plans, and contributing to building data repositories. He has additional strong background in the development of research, guidance, and ethics related to global diseases affecting resource-poor settings and orphan diseases while leading and/or contributing to challenging projects. He brings substantial experience (UNAIDS, WHO, and others, including local organizations and industry) in developing health-related research projects, collaborative engagements, regulatory and policy outreach, and education and training in Europe, Africa, Asia, the Americas, and Eastern Europe & Central Asia.



**Professor Perihan Elif Ekmekçi** (M.D., Ph.D.) (Co-Chairperson) is Head of History of Medicine and Ethics Department and Deputy Dean of TOBB ETU Medical School in Ankara, Turkey. She was a Fogarty Fellow in 2013-2014 and she had her master's certificate on research ethics through this fellowship from Harvard University. Dr Ekmekci was a research fellow in Imperial College Tanaka Business school in 2006. In 2015 she was elected as the first Turkish fellow for Western Institutional Review Board Research Ethics Program. She served as the Head of EU Relations Department of Ministry of Health Turkey (2007-2016). She was the Turkish representative for the European Center for Disease Control Advisory Board. She serves as the Head of the Institutional Review Board (IRB) of TOBB ETU and an Executive Committee Member of the Health Sciences Institute of TOBB ETU. She is a member of Open Science Committee of TOBB ETU. She established the UNESCO Bioethics Unit at TOBB ETU in 2019. In 2020 she co-authored the book *Artificial Intelligence and Bioethics*. (Contributed to the development of the AIDV-WG. Not a member of the previous RDA COVID-19 Legal/Ethical Working Group.)

**Dr. Anne Cambon-Thomsen**, Emeritus Research Director at CNRS (French national centre for scientific research), working in a research Unit on epidemiology and population health at Inserm (National Institute for Health and Medical Research), and University of Toulouse III Paul Sabatier, Faculty of Medicine Toulouse, France. Specialist in human immunogenetics, with a master's in human biology, a degree in health ethics, she works in an interdisciplinary research team on bioethics and innovation in health, involving human and social sciences as well as health sciences <https://cerpop.inserm.fr/>; team BIOETHICS). After a postdoc in Copenhagen (Denmark), she directed two research units (INSERM and CNRS) in immunogenetics, immunopathology, and population genetics in Toulouse (France) between 1985 and 1997. She served three terms on the European Group on Ethics in Science and New Technologies (EGE).

**Mr. Alexander Bernier**, pursuing a Doctor of Juridical Science (S.J.D) at the University of Toronto, Faculty of Law, under the supervision of Professor Gillian Hadfield. His doctoral research concerns the effects of data regulation on self-assembled biomedical data commons, and law and economics perspectives on the governance and oversight thereof. Alexander obtained a Master of Laws from the University of Toronto Faculty of Law. At the Centre of Genomics and Policy, his research is primarily concerned with data protection law, open science, and research infrastructure. Alexander Bernier is a member of the European-Canadian Cancer Network's Internal Ethical Board and is the Ethics Officer of the Canadian Open Neuroscience Platform Ethics and Data Governance Committee.

**Dr. John Brian Pickering**, a Senior Research Fellow working in computer science, specifically interested in human interaction and acceptance of technology. After many years in industry, he came back to academia to increase my involvement with pure and applied research. He also chairs the Faculty Research Ethics Committee and sits on the DPIA panel for the University. He is currently working on a number of medical informatics projects focusing on respiratory diseases.

**Dr. Simon Parker**, Data Steward, Ethico-Legal Framework and Patient Participation, Data Hub Operations, Deutsches Krebsforschungszentrum, Heidelberg, Germany. Background in history (specialising on the Renaissance) and criminology with a focus on data in his PhD studies in quantitative criminology. Initially he worked with sensitive social and economic data and supported research projects to use data safely. He was previously Data Liaison Manager at Cancer Research UK where he oversaw the creation of infrastructure where patients' medical records could be stored and analysed.

**Prof. Claudia Bauzer Medeiros**, full professor of Computer Science at the Institute of Computing, University of Campinas (Unicamp), Brazil. She has received Brazilian and international awards for excellence in research, in teaching, and work in fostering the participation of women in computing. She is a member of the Brazilian Academy of Sciences. Claudia received the Anita Borg Agent of Change Award (2009), is a Commander of the Brazilian Order of Scientific Merit and a former Distinguished Speaker of the Association for Computing Machinery (USA). She is a Doctor Honoris Causa from Universidad Antenor Orrego, Peru (2007), and from Universite Paris-Dauphine, France (2015). Her research is centered on the management and analysis of scientific data, to face the challenges posed by large, real-world applications. In 1994, she created the Laboratory of Information Systems at Unicamp, one of first research laboratories in Brazil dedicated to solving multidisciplinary problems involving scientific data in interdisciplinary research. Since 2014, she has been responsible for the eScience program at FAPESP (the Sao Paulo Research Foundation, Brazil). She is a member of the RDA Council and a Member-at-Large of the Council of the ACM.

**Dr. Gustav Nilsson**, a researcher in medicine, and associate professor of neuroscience at Karolinska Institute in Sweden. At the Swedish National Data Service, he is an expert in medical data and coordinates the domain specialists. He also co-chairs the European Open Science Cloud (EOSC) Researcher Engagement & Adoption Task Force (REA TF) that focuses on engaging diverse research communities in order to increase their participation in EOSC.

**Case Statement**  
**EOSC Future & RDA**  
**Artificial Intelligence & Data Visitation Working Group**

**Dr. Dara Hallinan**, FIZ Karlsruhe - Leibniz Institute for Information Infrastructure, a legal academic focused on the interaction between law, new technologies – particularly ICT and biotech – and society. He wrote his PhD at the Vrije Universiteit Brussel on the better regulation of genetic privacy in biobanks and genomic research through data protection law. He is also programme director for the annual Computers, Privacy and Data Protection Conference.

**Dr. Diana Dimitrova**, a post-doc researcher at FIZ Karlsruhe – Leibniz Institute for Information Infrastructures. She completed her PhD at VUB/LSTS in August 2021. In her PhD she studied the rights of access and rectification in the EU Area of Freedom, Security and Justice, with a focus on the border control context. Next to performing research on different data protection topics, she is also one of the core members of the CPDP Programming Committee and one of the editors of the Data Protection Insider, a biweekly newsletter issued by Lexxion. Before joining FIZ-Karlsruhe and commencing her PhD at VUB, Diana was a researcher at CiTiP, KU Leuven and a trainee at the EDPS.

**Mrs. Stéphanie Rennes**, legal counsel and data policy specialist, Legal Department, Institut national de recherche pour l’agriculture, l’alimentation et l’environnement (INRAE).

## 9 MEMBERSHIP

The AIDV-WG will develop an open and inclusive membership across countries, regions, and sectors. The membership will engage data producers, users and stewards, across different disciplines and sciences as well as different stages of data processing and exchange. The AIDV-WG will reach out across the RDA international community on its neutral social platform and assist in furthering Open Science goals in data sharing, education, data management plans. We encourage members who are engaged in the certification of data repositories, disciplinary and interdisciplinary interoperability, and the development of new technological related to AI and DV.

## 10 ENGAGEMENT WITH EOSC-FUTURE, THE RDA COMMUNITY AND OTHER

The AIDV-WG will engage with relevant working groups in EOSC-Future and other EOSC projects as well as with the RDA Community, its interest groups and working groups, and other national, regional, and international organisations engaged in AI, DV, and Open Science.

## 11 TIMELINE

The AIDV-WG will complete its activities over an 18-month period, utilizing monthly web meetings, work streams, and focused discussions on its deliverables. The deliverables will all be due by the end of July 2023 to meet the final report deadline for EOSC-Future on August 31, 2023. The EOSC-Future project ends on September 30, 2023, as does its funding.

<b>Date</b>	<b>Activity</b>
<b>May 30 to August 1, 2022</b>	3 Preliminary Discussion Meetings following the awarding of the EOSC-Future Grant: 1 <sup>st</sup> Meeting and launch of the AIDV-WG Zoom meeting link: <a href="https://us02web.zoom.us/j/87444290886?pwd=UP0szA1FEo1Cqrtx90NzSSGq1YBw-V.1&amp;from=addon">https://us02web.zoom.us/j/87444290886?pwd=UP0szA1FEo1Cqrtx90NzSSGq1YBw-V.1&amp;from=addon</a> Passcode 123
<b>Monthly September 2022 to March 2023</b>	First Monday of each month from 14:00 to 15:00 CET (Brussels/Paris time) Zoom meeting link:



**Case Statement**  
**EOSC Future & RDA**  
**Artificial Intelligence & Data Visitation Working Group**

<b>Date</b>	<b>Activity</b>
	<a href="https://us02web.zoom.us/j/83954320774?pwd=1ASnb2ihA10xtyrEnXUIgCWRE09Nzpz.1&amp;from=addon">https://us02web.zoom.us/j/83954320774?pwd=1ASnb2ihA10xtyrEnXUIgCWRE09Nzpz.1&amp;from=addon</a> Passcode 123
<b>June-September 2022</b>	Development of the work streams for the deliverables Recruitment of the fellows for the mentorship programme
<b>September 2022</b>	Deliverable 1: Survey design and plan Deliverables 2, 3, 4, & 5: Research
<b>October-November 2022</b>	Deliverable 1: Survey launched, and responses collected Deliverables 2, 3, 4, & 5: Research
<b>December 1, 2022</b>	First assessment of the fellows
<b>December 15, 2022</b>	Midterm Report of the AIDV-WG
<b>December 2022 - January 2023</b>	Deliverable 1: Data analysis Deliverables 2, 3, 4, & 5: Outline of the plans
<b>March 1, 2023</b>	Second assessment of the fellows
<b>February-May 2023</b>	Deliverable 1: Article summarizing the results of the research Deliverables 2, 3, 4, & 5: Writing
<b>May-July 2023</b>	Deliverables 2, 3, 4, & 5: Publications
<b>August 15, 2023</b>	Final Report on the Working Party to the EOSC-Future Project
<b>September to March 29, 2023</b>	Dissemination of the results and
<b>December 2023 to March 2024</b>	Webinars and/or meetings on the AIDV-WG's results
<b>January 31, 2024</b>	Final assessment of the fellows
<b>March 29, 2024</b>	Final Report of the AIDV-WG