TWIN4DEM: Strengthening democratic resilience through Digital Twins

D1.3

Ethics Briefing Pack

Grant Agreement No	101178061	**Acronym**	TWIN4DEM
Full Title	Strengthening Democratic Resilience Through Digital Twins		
Start Date	1 January 2025	**Duration**	36 months
Project URL	https://twin4dem.eu/		
Deliverable	D1.3 Ethics Briefing Pack		
Work Package	WP1		
Deliverable Type		**Dissemination Level**	Public
Due Date of Deliverable	30/6/2025	**Actual Submission Date**	30/6/2025
Deliverable Identifier		**Deliverable Version**	1.0
Lead Beneficiary	ETICAS		
Authors	Gemma Galdon Clavell, Mariana Carvajal Sojo, Maite Del Amo (ETICAS)		
Co-authors			
Reviewers	Emma de Mooij, Clara Egger (EUR), Márton Gerő, Balázs Bekker (CSS), Giangiacomo Bravo (LNU)		
Security Assessment	☐ Passed	☐ Rejected	☒ Not Required
Status	☐ Draft	☒ Peer Reviewed	☒ Coordinator Accepted

Document history

Version	Date	Partner	Remarks
v.1	30.06.25	ETICAS	

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

1. Executive Summary

This Ethics Briefing Pack (D1.3) is part of WP1, and is the deliverable Task 1.3, which is dedicated to identifying and addressing the ethical, legal, and societal implications of TWIN4DEM. It complements the Data Management Plan (D1.2) by focusing not just on how data is handled, but also on human participation, privacy, fairness, safety, and the responsible use of technology in the project. This document is designed to guide the consortium in ethics and integrity in every stage of the project.

The document outlines the core obligations of TWIN4DEM under Horizon Europe, the GDPR, and the AI Act. It explains how these frameworks apply to the project's work on digital twins and democratic resilience, particularly given the involvement of human participants and the use of sensitive political data. It provides clear direction on complying with legal requirements while safeguarding fundamental rights. A central focus is on practical measures: the Ethics Briefing Pack offers guidance on recruitment strategies, informed consent, and ethics approvals for research involving human participants. It also addresses how data will be collected, minimised, anonymised, and securely managed, with special attention to public data sources such as parliamentary records and official statements. This ensures the project's data practices are both lawful and respectful of individual rights.

Finally, the Ethics Briefing Pack promotes fairness and transparency in authorship, dissemination, and collaboration. It provides principles for recognising meaningful contributions to research outputs and highlights the importance of responsible communication of results. Overall, the Ethics Briefing Pack aims to help partners apply these principles consistently and with care throughout the project's lifecycle.

Table of Contents

List of Tables

ABBREVIATIONS

Abbreviation	Description
AI	Artificial Intelligence
API	Application Programming Interface
CSO	Civil Society Organisation
CSS	Computational Social Sciences
DMP	Data Management Plan
DT	Digital Twin
EC	European Commission
ECHR	European Convention on Human Rights
EU	European Union
GDPR	General Data Protection Regulation
GPAI	General-purpose AI
HE	Horizon Europe
MEPs	Members of the European Parliament
PEO	Project Ethics Officer
SSH	Social Sciences and Humanities
ToS	Terms of Service
UN	United Nations
WP	Work Package

1. Introduction

The TWIN4DEM project is committed to upholding the highest standards of research ethics and data protection across all its activities. Given the interdisciplinary nature of the consortium and the diverse types of data involved, the Ethics Briefing Pack aims to provide a clear framework for partners to ensure that ethical, legal, and responsible research principles are consistently applied throughout the project. This document outlines the main ethical requirements, procedures, and best practices to guide all activities involving human participants, data collection, data processing, and dissemination.

Research within TWIN4DEM is guided by a robust ethical framework grounded in European Union regulations, including the General Data Protection Regulation (GDPR), Horizon Europe (HE) ethical standards, and regional and international human rights instruments such as the EU Charter of Fundamental Rights, the European Convention on Human Rights (ECHR), European Commission (EC) guidelines and the ALLEA Code of Conduct for Research Integrity. These standards collectively establish the baseline for safeguarding human dignity, autonomy, and privacy in research, while promoting principles such as transparency, accountability, and fairness throughout the project.

TWIN4DEM aims to enhance the use of Computational Social Sciences (CSS) in democracy research by introducing Digital Twin technologies to analyse democratic resilience in Czechia, France, Hungary, and the Netherlands. This approach seeks to capture the complexity of political systems through participatory and co-construction methodologies, reflecting the ethical, legal, and societal implications of digital transformation in governance. Given the sensitivity of this research, ethical considerations and risk mitigation are addressed in this deliverable.

This Ethics Briefing Pack (D1.3) is part of WP1, and is the result of Task 1.3, which is dedicated to identifying and addressing the ethical, legal, and societal implications of TWIN4DEM. It is closely aligned with the Data Management Plan (DMP, D1.2), which addresses data collection and processing. However, the EBP focuses on human participation and broader ethical issues such as safety, fairness, and inclusivity. The relationship between the Ethics Briefing Pack and WP6 is also important, as the latter focuses on ethics-driven CSS, which will be materialized in a FAIR methodological toolbox (D6.1). Therefore, the Ethics Briefing Pack plays a key role in clarifying the ethical principles that will guide the research.

This deliverable is structured as follows: it begins by explaining the formal ethical obligations under Horizon Europe, the GDPR, and the AI Act. It then outlines the ethical considerations that guide all activities involving human participants, including the recruitment strategy, informed consent, and ethics approvals. The following section addresses the project's data sources, including the legal basis, the identification of sources, strategies for protecting data privacy (such as minimization and anonymization), and the grounds supporting access to personal data. It also details the ethical safeguards for the use of parliamentary, legislative, and official data, since this is public data and subject to different regulations. Additionally, it covers concerns related to data scraping and compliance with the Terms of Service (ToS) of digital platforms.

In line with the project's commitment to transparency, fairness, and consistency, the deliverable then includes authorship guidelines to ensure a shared understanding of what constitutes meaningful contributions and how they should be recognized. It sets out guiding principles for responsibility and accountability in authorship, along with criteria for determining direct and significant contributions to the research outputs, which will inform the order of authors. Lastly, the document addresses dissemination strategies, a key component of TWIN4DEM, acknowledging the responsibilities involved in sharing information derived from human participants.

This Ethics Briefing Pack is designed as a living document that will be updated as the project evolves, ensuring that TWIN4DEM remains responsive to emerging ethical challenges and aligned with best practices in research integrity.

2. Ethical guidelines, directives and legal frameworks

TWIN4DEM operates within the formal ethical obligations established by the European Commission for Horizon Europe-funded research and innovation activities in line with Regulation (EU) 2021/695[1]. At its core, this framework draws on fundamental human rights instruments such as the EU Charter of Fundamental Rights[2], the European Convention on Human Rights[3], the UN Universal Declaration of Human Rights[4], among others. Similarly, it establishes research ethics codes drawing inspiration from the Nuremberg Code[5], the Declaration of Helsinki[6] and the Belmont Report[7]. The combination of these sources, ultimately enshrined in Horizon Europe, build a foundational basis for ethical guidance in research, they reflect the highest standards for human treatment across disciplines.

This section outlines the regulatory framework that ensures compliance with ethical standards guiding the TWIN4DEM project. It integrates the requirements of Horizon Europe, the General Data Protection Regulation (GDPR), and the AI Act, each of which contributes a distinct yet complementary perspective on the ethical considerations. On one hand, Horizon Europe sets the foundation by emphasizing the protection of human participants and the preservation of fundamental values such as dignity, autonomy, and fairness. The GDPR, by contrast, focuses specifically on data protection, addressing the rights of individuals whose data is collected, processed, and stored. It establishes strict principles around data minimization, purpose limitation, and transparency, ensuring that the personal data of participants is handled responsibly and securely. Finally, the AI Act introduces a risk-based approach, classifying high-risk systems, such as those potentially used in TWIN4DEM, as requiring heightened safeguards.

2.1 Horizon Europe: Regulation (EU) 2021/695 and European Commission Guidelines

TWIN4DEM is bound by the ethical and legal obligations outlined in Regulation EU2021/695, which established Horizon Europe, the EU's key funding programme for

[1] European Union. 2021. "Regulation - 2021/695 - EN - EUR-Lex." Europa.eu. 2021. https://eur-lex.europa.eu/eli/reg/2021/695/oj/eng.

[2] European Convention. (2000). Charter of Fundamental Rights of the European Union . Official Journal of the European Communities. https://www.europarl.europa.eu/charter/pdf/text_en.pdf

[3] European Convention. (2000). Charter of Fundamental Rights of the European Union . Official Journal of the European Communities. https://www.europarl.europa.eu/charter/pdf/text_en.pdf

[4] United Nations. (1948, December 10). Universal declaration of human rights. United Nations. https://www.un.org/en/about-us/universal-declaration-of-human-rights

[5] British medical journal . (1996). Nuremberg Code . BRITISH MEDICAL JOURNAL.https://media.tghn.org/medialibrary/2011/04/BMJ_No_7070_Volume_313_The_Nuremberg_Code.pdf

[6] World Medical Association. (2013, October). WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Participants. Wma.net; The World Medical Association. https://www.wma.net/es/policies-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/

[7] Office for Human Research Protections. (1979). The Belmont Report. U.S. Department of Health and Human Services. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html

research and innovation. Article 19 of this Regulation is particularly relevant, as it sets out the need for all research and innovation activities to comply with fundamental ethical principles and applicable Union, national, and international law. It specifies ethics procedures that funded projects must follow. For instance, ethics assessments to identify areas of complex issues, monitoring throughout the life cycle of the project, and its subsequent ethics checks carried out by the Commission. Horizon Europe positions ethics as a proactive responsibility that must be embedded in project design, implementation, and dissemination. This is particularly relevant for projects like TWIN4DEM that explore themes of democracy and political participation through technological tools. Therefore, it is crucial to adopt a proactive ethical approach to protect fundamental rights, protect participants, and promote public trust in innovation.

To guide this work, TWIN4DEM draws on a range of European Commission guidelines and ethical frameworks. The European Commission's Ethics for Social Science and Humanities[8] (SSH) research guide offers practical direction for researchers working in contexts where ethical questions are central. Importantly, it reminds all partners that ethical obligations extend to any research that impacts people's identities, rights, or well-being irrespective of the discipline.

In line with these guidelines, researchers working under Horizon Europe are expected to:

- Respect human dignity and integrity
- Ensure honesty and transparency towards research participants
- Promote individual autonomy and obtain free and informed consent
- Take special care to protect vulnerable groups
- Safeguard privacy and confidentiality
- Advance justice and inclusiveness across research processes
- Minimise potential harm and maximise social benefit
- Share benefits with disadvantaged populations
- Protect the environment and future generations where applicable.

These principles are enforceable standards that EU-funded projects must address at the design, implementation, and dissemination stages of each activity.

In addition, the European Commission's Guidelines on Ethics and Data Protection in Research[9] further clarify expectations when research involves the use of personal data. The Guidelines underscore that when data processing activities entail risks to the rights

[8] European Commission. 2021. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-in-social-science-and-humanities_he_en.pdf#page=4.64.

[9] European Commission. 2021. "Ethics and Data Protection." https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-and-data-protection_he_en.pdf.

and freedoms of individuals, these must be clearly communicated during the informed consent process. Moreover, consent should not only be informed, but also meaningful, voluntary, and revocable.

Importantly, Horizon Europe recognises that ethical concerns extend beyond direct human participation. Activities that involve the collection, classification, or algorithmic treatment of personal data can raise ethical issues even when no person is physically involved. For instance, inferring political affiliation from digital behaviour, even from public sources, must be evaluated not only for legal compliance, but for its broader ethical implications. Ultimately, the activities must follow clear and rigorous standards to ensure that, even in heightened risk scenarios, the project is compliant.

To reinforce these commitments, TWIN4DEM also relies on the European Code of Conduct for Research Integrity, developed by ALLEA[10] and endorsed by the European Commission. The Code outlines four key values: reliability, honesty, respect, and accountability. Together, these values guide TWIN4DEM's internal protocols for research design, authorship, peer review, data handling, and public engagement. These frameworks shape a robust and applicable ethical infrastructure for TWIN4DEM. Hence, ethics within the consortium, ethics is treated as a measure to ensure that all research process and outcomes uphold democratic values, protect human dignity and fundamental rights, and serve the public good.

2.2 General Data Protection Regulation

The General Data Protection Regulation (GDPR)[11] is the main legal framework governing the processing of personal data in the European Union. It applies to any activity that involves collecting, storing, processing, or sharing information about identifiable individuals, regardless of whether that data is collected directly by the project or reused from external sources. As a Regulation rather than a Directive, the GDPR has direct effect in all Member States, imposes binding obligations on all Horizon Europe-funded activities, including those undertaken by the TWIN4DEM consortium. The GDPR frames the rights of data subjects and the responsibilities of controllers and processors to ensure that personal data is handled lawfully, ethically, and transparently.

In the context of TWIN4DEM, these obligations are particularly relevant where human participants are involved in activities such as surveys, interviews, focus groups or processing of online data. In these cases, the consortium must ensure compliance with the GDPR when handling any information that relates to an identified or identifiable

[10] ALLEA. 2023. "The European Code of Conduct for Research Integrity REVISED EDITION 2023," June. https://doi.org/10.26356/ECOC.

[11] European Union, "General Data Protection Regulation," Europa.eu, April 27, 2016, https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng.

individual (Article 4(1)). This includes not only obvious identifiers like names or email addresses, but also online identifiers, geolocation data, demographic profiles, and metadata that can, directly or indirectly, be linked to a person.

Furthermore, TWIN4DEM must pay close attention to the processing of special categories of personal data as defined in Article 9(1), which includes political opinions. This is especially relevant to the project, as it aims to study democratic processes and resilience, potentially through questions related to civic participation, trust, or voting behaviour. Article 9(2) provides a limited set of legal bases under which such data can be processed, most relevantly when data subjects give explicit consent (Article 9(2)(a)), or when processing is necessary for scientific research purposes (Article 9(2)(j)), subject to appropriate safeguards under Article 89(1).

In addition, the consortium must uphold the six core data protection principles enshrined in Article 5(1) GDPR:

- **Lawfulness, fairness and transparency** (Article 5 (1) a), according to Article 5 (1) a: "personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency')." Therefore, partners must process personal data lawfully, fairly and in a transparent manner in relation to the data subjects who will take part the pilots developed in WP7.
- **Purpose limitation** (Article 5(1)(b)), according to this Article, personal data shall additionally be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation'). This principle compels partners to collect data only for clearly stated and justified purposes. Ethically, this ensures that data subjects retain a degree of control and understanding over how their data contributes to the project.
- **Data minimisation** (Article 5(1)(c) GDPR) states that no data should be collected if they are not strictly necessary for the declared purposes of the processing. In other words, if the utility of a piece of data is unclear, it should not be collected. Evidently, this requires a contextual judgement that takes into consideration the purpose of the processing and the suitableness of data in order to achieve it.
- **Accuracy** (article 5(1)(d) and 16) is essential in AI development. Incorrect data can produce harmful outcomes, especially when automated systems are involved. Ethical research therefore demands continuous monitoring and correction of inaccurate or outdated data. All data subjects whose personal data is managed by the Project have the right to request that Project partners erase or rectify without delay erroneous data that relates to them. The consortium must

take every reasonable step to update or remove data that is inaccurate or incomplete. This principle is connected to the right to rectification contained in article 16 GDPR, which grants data subjects the right to demand the revision of the personal data that is not accurate. The importance of this principle arises from the potential damage that can be caused to a data subject if inaccurate data is associated to them.

- **Storage limitation** (article 5(1)(e)) entails that all partners must delete personal data when they no longer need it and as it is said above, the concrete retention period/s to be established are not fixed by law. Therefore, it must be determined on a case-by-case basis in attention to the nature of the processing and its purposes. The retention period needs to be justified on the grounds of its utility. No data can be held if they do not serve the purposes for which they were collected in the first place.
- **Integrity and confidentiality** (article 5(1)(f)) require that data is securely protected from breaches or misuse. Beyond technical compliance, this principle reflects the ethical obligation to prevent harm to individuals whose data is entrusted to the project. All partners must keep personal data safe, secure and protected by using appropriate technical and/or organisational measures.
- **Pseudonymisation and anonymisation** of data encourages partners to place appropriate safeguards on data. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner.
- **Processing of special categories of data** (Article 9), such as health information or biometric data is in principle prohibited, unless under certain conditions, such as obtaining explicit consent and respecting heightened safeguards. This is particularly relevant in ACHILLES pilots, where vulnerable groups may be involved or where inferences may be drawn from seemingly neutral but special personal data.

To further comply with GDPR obligations, TWIN4DEM must also consider the principles of data protection by design and by default (Article 25), ensuring that privacy measures are embedded into the technical development of digital tools and administrative workflows from the outset. This proactive approach is particularly important when introducing new processing activities involving sensitive data or large-scale datasets, as it reduces the risk of data breaches and ensures compliance from the beginning. In cases where such high-risk processing is involved, a Data Protection Impact Assessment (DPIA) may be required under Article 35, to evaluate and mitigate potential privacy risks.

Additionally, Article 89(1) of the GDPR requires that all research-related data processing be accompanied by suitable safeguards to protect the rights and freedoms of data subjects. These safeguards may include:

- **Pseudonymisation**, when identifiers are replaced with coded references;
- **Anonymisation**, where identification is no longer possible (with caution, given the risk of re-identification);
- **Restricted access controls**, applied through role-based permissions;
- **Encryption**, where data are stored or transmitted.

TWIN4DEM partners must also uphold the data subject rights defined in Chapter III of the GDPR, including the rights to access (Art. 15), rectification (Art. 16), erasure (Art. 17), restriction of processing (Art. 18), and objection (Art. 21). These rights should be communicated to participants as part of the consent process, and mechanisms must be in place for partners to respond to any such request.

Finally, it is important to note that even when personal data is drawn from publicly accessible sources, such as public statements by political figures, this data remains subject to GDPR unless it has been genuinely anonymized. This distinction is critical, as public availability does not eliminate the need for data protection safeguards. Moreover, Recital 33 of the GDPR acknowledges that in some research contexts, it may not be possible to fully identify the purposes of data processing at the time of collection. This allows for more flexible consent models, where participants can consent to specific aspects of research while retaining the right to withdraw or limit their consent as the project evolves. This reinforces the principle of consent as a dynamic and ongoing process, rather than a one-time transaction.

2.3 Artificial Intelligence Act

The Artificial Intelligence Act (AI Act), formally adopted by the European Union in 2024, establishes a harmonised legal framework for the development, deployment, and use of artificial intelligence in the EU. Although the Act will apply fully from 2 August 2026, several provisions become enforceable earlier, which has implications for the TWIN4DEM project timeline and outputs. The TWIN4DEM consortium adopts a forward-looking approach to demonstrate its commitment to align the project's activities with the obligations in the AI Act. This ensures the AI systems designed, developed, and used in the project (even afterwards) are and will remain sustainable.

Under the AI Act, according to Article 3(1) any system based on machine learning, statistical inference, logic-based or knowledge-based systems falls within the Act's scope. This broad definition means that many of the data processing and simulation tools in TWIN4DEM qualify as AI systems, even if they are not marketed as such. This carries two important implications: First, some use cases in TWIN4DEM may fall under the "high-risk AI" classification, particularly those that analyse or predict political orientation or opinions. These applications are explicitly listed in Annex III, point 11 of the AI Act as high-risk when they infer sensitive attributes.

If classified as high-risk, the following obligations apply under Title III, Chapter 2 of the AI Act:

- **Risk management system** (Article 9): The system must undergo continuous identification, analysis, and mitigation of risks throughout its lifecycle.
- **Data governance and data quality requirements** (Article 10): Training, validation, and testing data must be relevant, representative, free of errors, and complete to avoid bias and misclassification.
- **Technical documentation and record-keeping** (Article 11): Documentation must be detailed and kept up to date, enabling authorities to assess compliance.
- **Transparency and provision of information to users** (Article 13): End-users must be informed of the AI system's capabilities, limitations, and appropriate use.
- **Human oversight** (Article 14): Adequate measures must be in place to allow human intervention, oversight, and, where necessary, override of the system.
- **Accuracy, robustness, and cybersecurity** (Article 15): Systems must meet high standards for technical performance and be resilient to errors and malicious exploitation.

In addition, Article 52 highlights the obligation of systematic risk evaluations, explicitly including data collection and processing. These evaluations must account for the protection of fundamental rights and incorporate appropriate mitigation measures. Articles 9 and 52 are thus closely linked in establishing a proactive and continuous risk management process. Furthermore, user rights and anti-discrimination safeguards are reinforced by Article 23, which emphasises transparency, safety, and the obligation to carry out conformity assessments, ensuring that technical documentation remains relevant and up to date. Should new risks arise, Article 23 also mandates active cooperation with the competent authorities.

Second, for AI systems that do not qualify as high-risk but are considered general-purpose AI (GPAI), Chapter V of the AI Act outlines specific responsibilities. These include obligations for transparency, appropriate labelling, and traceability. In TWIN4DEM, if models trained on open data or archival texts are repurposed across use cases or domains, these provisions will still apply.

Additionally, Article 25 is particularly relevant regarding the management of data sources. It establishes that any distributor, importer, or third-party modifying AI systems (including their data sources) must comply with the regulation requirements. This includes undergoing conformity assessments and maintaining appropriate documentation. Importantly, Article 25 extends compliance obligations to agreements between providers and third parties, also protecting intellectual property rights.

Given that the project will be developing population models, policy simulations, and data-driven insights into democratic dynamics, all algorithmic components must be designed with ethical safeguards and legal compliance in mind. The project's reliance on political, institutional, and social datasets introduces a heightened obligation to ensure the integrity and social impact of its tools. Finally, the AI Act underlines the role of public-interest research (Recital 14 and Article 2(5)) and provides certain flexibilities for scientific institutions. However, these exceptions do not waive the obligation to ensure that systems are secure, documented, explainable, and aligned with fundamental rights.

3. Humans

This section outlines the ethical and legal considerations that guide all activities involving human participants in TWIN4DEM. These activities, including interviews, and focus groups, are integral to the project's mission of enhancing democratic resilience through digital twin technologies. They aim to capture diverse perspectives, provide critical insights into the project's design and implementation, and assess the broader social implications of the technologies being developed. In particular, these methods are essential for understanding the public's perception of TWIN4DEM, identifying potential risks, and evaluating the project's societal impact.

Focus groups will be conducted at various stages, for instance, WP7 will organise focus groups for the country case consortium partners. By engaging directly with participants, TWIN4DEM seeks to create a more inclusive and societally relevant digital twin prototype , while also contributing to the evidence base on democratic backsliding and governance challenges in Europe.

While the project's broader legal and regulatory obligations are covered in the preceding sections (Horizon Europe, GDPR, and the AI Act), the following focuses on their operationalisation in the context of recruitment, informed consent, and ethics approvals. These measures aim to ensure that the rights, dignity, and safety of all individuals involved are respected at every stage of the research process.

Scope of human participation in TWIN4DEM

TWIN4DEM engages human participants primarily through several actities focus groups. Based on the current work plan, the activities directly involving human subjects are coordinated by Democracy International and includes the following beneficiary entities: ICL, GESIS, FBK, LNU, UBB, CSS, CUNI, ETICAS, DI and DBC under specific WP2, WP4, WP5 and WP7. These activities are bound by the appropriate ethical considerations defined in Horizon Europe and the GDPR, for research involving human participants. Therefore, it must adhere proper recruitment methods, material and values such as informed consent (Art 3.59, AI Act), communication and voluntariness (Art. 7, GDPR). Hence, all participation must be fully voluntary, and consent must be informed, explicit, and reversible.

In addition, the TWIN4DEM Training Camp (Task 6.3) involves the recruitment and engagement of junior and early-career researchers for capacity-building purposes. While this activity does not involve research data collection or human subjects contributing to scientific outputs, it will follow best practices in data protection, inclusivity, and fairness in participant selection. Personal data collected for registration and participation will be handled in line with GDPR principles, with transparency about its use and storage. Both the recruitment of participants and the organisation of activities have been and will continue to be accompanied by ETICAS through the

review of appropriate consent forms and recruitment strategies, particularly when engaging external participants outside of the consortium.

3.1 Recruitment strategy

TWIN4DEM's recruitment strategy for the consent groups is grounded in a commitment to fairness, transparency, and inclusivity, in line with the ethical principles outlined in the Horizon Europe framework, the GDPR, and international research ethics guidelines like the ALLEA Code of Conduct for Research Integrity. Given the project's emphasis on democratic resilience, the approach to recruitment aims to ensure that a wide range of voices is included while safeguarding participants' rights and data privacy.

There has been set different recruitment methods depending on the target audience and focus groups objectives. The possible methods are the following:

1. Targeted outreach: specific dissemination through partner organizations. Use of Democracy International's network. Partner institutions can propose contacts based on their local expertise.
2. Personalized invitations to key stakeholders, such as policymakers and leading civil society.
3. Referrals and snowball sampling: participants may suggest new candidates.

Further, online focus groups have open calls and public invitations via social media or different kinds of social networks such as academic or professional.

The recruitment process for TWIN4DEM has been designed to promote fairness, inclusiveness, and diversity, in line with EU research ethics guidelines for social science. No participant will be excluded on the basis of protected characteristics such as age, gender, ethnicity, disability, or political opinion without a justified reason connected to the research aims. At the same time, positive measures will be applied to ensure balanced and meaningful representation across gender, socioeconomic background, and other relevant dimensions, in order to reflect the diversity of the societies being studied.

This approach aligns with the principles of justice and inclusiveness outlined in Horizon Europe and EU research ethics guidelines, which encourage the fair representation of different social groups in social science research. However, certain exclusion criteria are necessary to safeguard the integrity of the research. For example, individuals under the age of 18 will not be eligible to participate, individuals with active electoral campaigning goals or those intending to use the focus group platform for propaganda, lobbying, or extremist messaging will be excluded. These criteria aim to prevent the misuse of the research setting and to ensure that focus group discussions remain constructive and aligned with the project's objectives.

Transparency is a cornerstone of TWIN4DEM's recruitment process. All recruitment materials will be clear, concise, and accessible, avoiding technical jargon and providing potential participants with a straightforward overview of the project's aims, the expected time commitment, and their rights. This includes a plain-language summary of the project, a description of the recruitment process, and contact details for any questions. Importantly, informed consent must be obtained before participation, with participants clearly informed of their right to withdraw at any time, without penalty, as required by Article 7 of the GDPR. The consent process must also address the specific nature of data processing, including any data sharing, publication, and retention plans, in line with Articles 13 and 14 of the GDPR.

To ensure that participation is genuinely voluntary, all participants will have at least one week to consider their involvement before signing a consent form. They will be explicitly informed of their right to withdraw at any point, without needing to provide a reason and without facing any negative consequences. If a participant chooses to withdraw, their data will be deleted upon request, unless they consent to anonymized retention for ongoing research purposes, ensuring compliance with the GDPR's requirements for data erasure (Article 17).

The recruitment strategy adheres to the principle of data minimization (Article 5(1)(c) GDPR), collecting only essential personal data necessary for the study's aims. All data will be securely stored, with appropriate access controls and encryption to prevent unauthorized access. This includes explicit consent for any recordings and clear communication regarding the data retention period. Focus group discussions will be conducted in secure settings to protect sensitive information and prevent unauthorized disclosures, aligning with the GDPR's principles of data integrity and confidentiality (Article 5(1)(f)). Finally, the recruitment process is managed by designated leads within the consortium and advised by ETICAS to ensure consistency, transparency, and adherence to ethical guidelines.

3.2 Informed consent

All participants will receive comprehensive information about the study's objectives, how their data will be used, their rights as data subjects (including the right to withdraw as established in Article 7 of the GDPR), and relevant contact points for any questions or concerns. In line with the GDPR, informed consent must be a freely given, specific, informed, and unambiguous indication of the participant's agreement to the processing of their personal data (Recital 32). Consent must be explicit for each intended purpose (Article 6) and participants should be made aware that they have the option to consent to only certain aspects or stages of the research, as highlighted in Recital 33.

To ensure compliance, researchers must be able to demonstrate that valid consent has been obtained (Article 7), including maintaining clear records of when and how consent

was provided. Therefore, TWIN4DEM will implement consent procedures that are transparent, accessible, and tailored to the specific nature of each research activity, using clearly written consent forms that are easy to understand and designed to support fully informed decision-making by participants.

Clear, concise, and accessible communication with participants is a core requirement under the GDPR, ensuring that all individuals can make informed decisions about their involvement. This includes providing a plain-language summary of the project, clearly outlining eligibility criteria and the recruitment process, detailing what participants can expect during their involvement, and including comprehensive informed consent forms that must be signed before participation. These forms should also provide contact details for the recruitment lead, allowing participants to ask questions or raise concerns at any stage.

To facilitate this process, a standardized template for informed consent, covering both written and oral formats, is included in Annex I of this deliverable. This template will be adapted as needed to align with the specific methods of data collection used in the TWIN4DEM project.

3.3 Ethics approvals

Ethics approvals serve as a formal mechanism to ensure that all planned activities involving human participants are assessed for potential risks and align with established ethical principles, legal obligations, and best practices. This includes evaluating the methods, objectives, and anticipated impacts of the research to prevent harm and uphold participant rights. They involve a review of study protocols, consent procedures, data protection safeguards, and risk mitigation strategies. Their purpose is to protect the rights, dignity, and welfare of participants, ensure responsible data handling, and promote research integrity. Ethics approvals are particularly important in TWIN4DEM, given the project's work with human participants, personal data, and sensitive topics such as political opinions.

Each partner involved in activities that include human participants or personal data processing is responsible for seeking ethics approval from their own institutional ethics boards or equivalent bodies before starting data collection. To support consistency, ETICAS (as PEO) will provide oversight by reviewing deliverables and planned activities to ensure that no required ethics approval is overlooked. Ethics approvals are necessary whenever a partner plans to collect data from human subjects, process identifiable personal data, or introduce significant changes to the methods that could pose ethical or legal risks.

Partners are encouraged to rely on informed consent templates, participant information sheets, and communication formats being developed in WP7 to support ethics submissions and interactions with participants. Where no institutional ethics process

exists, partners will liaise with ETICAS to define suitable alternative review arrangements. In cases where ethics approval is not granted, partners are required to thoroughly document the application process, including any feedback or decisions from the ethics committees. These records must be included as annexes in relevant project deliverables to maintain transparency and traceability, and to demonstrate the consortium's proactive approach to ethical oversight. Additionally, the consortium will report on the timing and scope of ethics approvals during periodic project reviews to ensure ongoing compliance.

4. Data Sources

This section provides an overview of the project's data sources and the associated legal requirements under Horizon Europe, the AI Act, and the GDPR. It will address critical aspects such as the identification of sources, data minimization strategies, anonymization approaches, and the legal basis for data access. These subsections will outline the methods and safeguards used in TWIN4DEM to ensure full compliance with European Commission regulations and robust data protection practices.

4.1 Horizon Europe and European Commission Guidelines

TWIN4DEM operates under the ethical and legal standards established by the European Commission for Horizon Europe research activities. According to the EC Guidelines on Ethics in Social Science and Humanities, data collection and analysis (even from open sources) must be approached with the understanding that large-scale datasets relating to social phenomena ultimately concern individuals. Researchers therefore bear a direct responsibility to prevent harm, safeguard rights, and minimise risks of misuse or unintended re-identification.

The European Commission has clarified that the availability of data in public spaces, particularly social media, does not equate to unconstrained re-use for research purposes. The Guidelines specifically warn against the assumption that publicly accessible online data may be used freely without considering ethical and legal limitations. In cases where researchers are collecting social media content or republishing user-generated data, the EC recommends three practical safeguards:

- paraphrase all data that will be republished (to prevent others being led to the individual's online profile),
- seek informed consent from people whose data you intend to use in its original form in research outputs, or
- consider a more traditional research approach that better ensures consent and confidentiality.

The Guidelines further emphasize that researchers must review and respect the terms of service of the platforms from which data are collected. Even if user content is publicly visible, scraping or automated collection may violate contractual obligations established by those platforms and potentially expose the project to legal risk.

In parallel, the EC's Guidelines on Ethics and Data Protection reaffirm that data protection is a cornerstone of research ethics. It is not merely a regulatory requirement but a reflection of fundamental rights under Article 8 of the EU Charter of Fundamental

Rights and Article 16 of the Treaty on the Functioning of the European Union (TFEU)[12]. This ethical framing reinforces the need for responsible collection and processing of data, particularly when it involves political opinions or any information that could be linked to identifiable persons, even indirectly.

This approach is especially relevant for TWIN4DEM, where public political content, parliamentary data, and social media outputs may be used in the CSS in democratic research. While some of these sources are manifestly public, the consortium commits to high standards of ethical scrutiny, informed by the broader legal framework and specific requirements of Horizon Europe, to ensure that all data activities respect the dignity, autonomy, and privacy of individuals.

4.2 General Data Protection Regulation

4.1.1 Identification of sources

Each country case partner will map and document the data sources used within their respective tasks. These sources include, but are not limited to, open-access repositories, institutional databases, legal archives, social media platforms, and public records. This approach aligns with the European Code of Conduct for Research Integrity (ALLEA, 2023), which emphasizes the importance of data transparency and accountability in research. Specifically, the ALLEA Code incorporates the FAIR principles (Findable, Accessible, Interoperable, and Reusable) as foundational values for data management, ensuring that data is "as open as possible and as closed as necessary." These principles emphasize the critical balance between open science and the protection of personal data.

Moreover, to comply with these principles, TWIN4DEM must ensure transparency regarding the access to and use of personal data, code, software, and other research materials. The identification of sources also serves as a foundation for ensuring that all data used in the project is lawfully collected, appropriately licensed, and ethically sound. Below, the different data sources used in TWIN4DEM are outlined based on the project's structure, however, a more detailed overview of data is specified in the Data Mangement Plan (D1.2):

- **Task 3.1 (CUNI)**: Data collection through web scraping and content analysis, including public reports, social media, and existing databases.
- **Tasks 3.2 and 3.3 (FBK)**: Use of experimental data from existing databases.
- **Task 3.4 (CSS)**: Collection of aggregated data from partners to form a new dataset.

[12] European Commission. 2012. " Treaty on the Functioning of the European Union ." 2012. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:12012E/TXT:en:PDF.

- **Tasks 4.1, 4.2, 4.3 (GESIS)**: Surveys that will use existing datasets, some of which are open access while others, such as EU_SILC, require formal registration and access approvals.
- **Task 7.1 (DI)**: Data from focus groups, collected via audio recordings, focused on Czechia, the Netherlands, France, and Hungary.
- **Tasks 5.2, 5.3, 5.4 (UBB)**: Simulations conducted using open data, expected to result in software outputs.

4.1.2 Data Minimisation Strategies

In alignment with the GDPR principle of data minimisation (Article 5(1)(c)), the consortium is committed to collecting only the data strictly necessary to achieve the project's objectives. Personal data that is not essential to the research questions will not be gathered or processed. This approach reduces the risk of re-identification and ensures compliance with the data protection by design and by default requirements set out in Article 25 of the GDPR.

Key minimization practices in TWIN4DEM include:

- Designing focus group engagement strategies to exclude unnecessary personal details or sensitive data not directly relevant to the research questions.
- Limiting metadata collection in scraping and aggregation activities to avoid storing superfluous or identifying attributes.
- Ensuring that aggregated or anonymised datasets do not derive from excessive raw data collection that could have been avoided.

4.1.3 Anonymization Strategies

Under the GDPR, anonymous data refers to information that does not relate to an identified or identifiable natural person, or to personal data that has been processed in such a way that the individual can no longer be identified by any means reasonably likely to be used.

Wherever personal data is involved, anonymization techniques will be applied to further reduce the risk of re-identification. These include pseudonymization, data aggregation, access controls, and technical safeguards as described in sections above. Article 25 of the GDPR strongly encourages the use of these techniques, while Recital 26 emphasizes that even pseudonymized data that can be attributed to a natural person by the use of additional information should still be considered personal data. The AI Act also supports these principles, requiring that datasets used to train AI systems follow sound data governance practices (Article 10).

4.1.4 Legal Basis for Access

All data collection involving personal data in TWIN4DEM must have a clearly established legal basis, typically informed consent for human participants or legitimate interest for public datasets. For archival materials and public speeches, partners must document access rights and reuse permissions to ensure compliance with GDPR and Horizon Europe requirements, as outlined in the D1.2 Data Management Plan.

The GDPR establishes the core principles for lawful data processing in Art. 5. Moreover, TWIN4DEM partners will ensure that data subjects' rights, such as the right to access (Art. 15), the right to rectification (Art. 16), and the right to be forgotten (Art. 17), are respected throughout the project. These rights also extend to data portability (Art. 20) and the right to object (Art. 21), which must be clearly communicated to participants during the consent process.

TWIN4DEM ensures legal access to data by collecting from official and publicly accessible sources, such as parliamentary records, EU decisions, and legislative texts. The project deliberately excludes non-public data, such as individual voting records or any restricted information. For social media data, the project recognises that the mere availability of information does not automatically grant permission to collect or reuse it. Therefore, before any data is gathered, the Terms of Service (ToS) of each platform must be carefully reviewed to confirm that collection methods comply with both the platform’s rules and legal requirements. Data scraping or automated data collection will not be conducted where it violates the ToS or contractual obligations. The project will give priority to platforms and data sources whose conditions align with GDPR and broader data protection principles.

4.1.5 Ethical safeguards for the use of parliamentary, legislative and official data

As outlined in the data summary table of D1.2, Data Management Plan, TWIN4DEM relies on publicly available data from parliamentary, legislative, and official sources, including speeches, roll-call votes, legislative texts, court rulings, and ministerial statements, primarily under the scope of WP3 and WP4.

Tasks under Work Package 3, --which are the creation of a corpus of multilingual textual data (T3.1), implementation of NLP tools to analyse political discourse (T3.2), implementation of NLP tools to classify executive aggrandisement in legal data (T3.3) and, development of an interlinked database on executive aggrandisement (T3.4), aim to improve the computational analysis methods. Moreover, these tasks identify cases of executive aggrandisement in adopted or debated legislation and analyse the responses to this phenomenon by political actors, courts, the EU, and civil society representatives.

Then, WP4 aims to provide integrated and synthetic data sets that include the data processed and compiled in WP3. This data will be further complemented bythe collection of survey data (T4.1), procuring elections, institutional, and official statistics data (T4.2) and an integration and updating of the data (T4.3) (see data flow in D1.2, Data Management Plan).

Both WPs are supported by the Data Management Plan (D1.2) and this Ethics Briefing Pack (D1.3), which collectively establish the framework for ethical data use, privacy protection, and compliance with EU regulations. These tasks will ultimately feed into the implementation and validation of digital twin prototypes, serving as the backbone for the project's computational social science methodologies.

While these datasets involve individuals acting in their public roles, they still fall under the protective scope of data protection laws when they include potentially sensitive information like political opinions. Under Recital 51 and Article 9 of the GDPR, data revealing political opinions is considered a "special category" and requires enhanced protection. However, Article 9(2)(e) allows for the processing of data "manifestly made public" by the data subject, such as parliamentary speeches and official statements, while Article 9(2)(j) permits processing for scientific research purposes, provided that appropriate safeguards are in place. These safeguards must ensure that the data processing is proportionate to the research aims and respects the fundamental rights of the data subjects, including their privacy and freedom of expression.

Moreover, the Open Data Directive (EU 2019/1024) encourages the reuse of public sector information, including parliamentary records, for research and innovation, emphasizing transparency and public access. While most of this data is not subject to copyright restrictions, the TWIN4DEM consortium is committed to following best practices in data ethics. This means using only what is necessary for research objectives, avoiding unnecessary linkage of public data to personal or unofficial data sources, and clearly documenting the purpose and scope of data collection.

4.1.6 Data scraping and compliance with platform Terms of Service

Web scraping, a technique for automatically collecting information from publicly accessible online sources such as news outlets, social media platforms, forums, and personal websites, can yield large volumes of personal data and therefore carries specific risks. As the European Data Protection Board has noted, indiscriminate scraping may affect many data subjects at once and create a pervasive sense of surveillance, potentially chilling freedom of expression and leading to self-censorship.

Under GDPR principles of lawfulness, fairness and transparency (Art. 5(1)(a)), purpose limitation (Art. 5(1)(b)), data minimisation (Art. 5(1)(c)), accuracy (Art. 5(1)(d)), storage

limitation (Art. 5(1)(e)), any personal data acquired through scraping must be processed lawfully, fairly, and with respect for the data subjects' rights. Equally important is adherence to the Terms of Service and usage policies of any platforms considered for data collection, whether TikTok, Facebook, Instagram, YouTube, or others. Automated collection is permitted only where expressly allowed by those terms, for example via official APIs with valid credentials. Collecting data in contravention of a platform's contractual conditions risks legal liability and undermines the ethical foundations of the research.

TWIN4DEM is committed to both data protection law and platform compliance. Data scraping efforts will be limited to those methods explicitly authorized by each service's Terms of Service. Where only manual or API-based access is permitted, automated scraping tools will be disabled in favour of approved channels. All scraping activities will be documented, justified by research necessity, and implemented in a manner that minimises unnecessary exposure of personal information.

Table 1 Mapping of main social medial platform terms of service

Platform	Terms of Service	Research-specific access	Risks
YouTube	YouTube's terms explicitly prohibit accessing its services via automated means, including scrapers, except for public search engines in compliance with its robots.txt file. Research access is generally limited to the YouTube Data API.	Research access is generally limited to the YouTube Data API.	Scraping outside the API is a direct violation of YouTube's ToS and can result in IP blocks, account suspensions, and potential legal action under laws like Directive 2013/40/EU on attacks against information systems.
TikTok	Section 5 of TikTok's terms prohibits the use of automated	Web scraping is accepted through the use of APIs.	TikTok actively enforces its anti-scraping policies, including litigation against large-

	scripts to collect information or otherwise interact with its services without prior authorization.		scale data scrapers. Use of data outside the official Research API is strictly prohibited. Violations may result in legal action and platform bans.
Instagram	Like Facebook, Instagram prohibits automated data collection without explicit permission. Research access is limited to the Instagram Graph API, which only provides data from accounts researchers own or have explicit consent to access.	No public research API. Researchers must apply for Instagram Graph API access, which returns only data from accounts they control or have been granted permission by.	Large-scale scraping of public profiles, posts, or stories without authorization violates the terms and can trigger account restrictions, data takedowns, or legal action, especially within the EU and US.
Facebook	Meta prohibits data scraping without explicit permission.	Its developer guidelines direct all data access through official APIs. Research access is typically granted through vetted partnerships like Social Science One or Data for Good, requiring contractual agreements.	Meta has a history of legal action against unauthorized data scraping, including lawsuits and account suspensions. Even projects developed for academic research (e.g., NYU Ad Observer) have faced blocks. Unauthorized scraping can lead to significant legal liabilities.

Data scraping involves complex legal considerations, particularly in the areas of personal data protection, intellectual property, and contract law, all of which are relevant to the TWIN4DEM project.

First, the collection of personal data through scraping must comply with the GDPR, as profile pictures, usernames, and other metadata can make the individuals who post or interact on social media identifiable, either directly or indirectly. This requires adherence to the core principles of data processing outlined in Article 5, including lawfulness, fairness, transparency, purpose limitation, and data minimisation. Furthermore, Article 25 mandates the implementation of technical and organizational measures to ensure data protection by design and by default, from the outset of data collection. Article 32 reinforces this requirement by obliging data controllers to implement measures to protect personal data against breaches or unauthorized access, even when such data is collected through automated methods like scraping.

Second, intellectual property rights also impose significant constraints. While not all social media content is subject to copyright, much of it is, as it reflects the creativity and original expression of its authors. Directive 2001/29/EC (the InfoSoc Directive) provides a legal framework for copyright protection in the EU, which is further reinforced by Directive 2019/790/EU, emphasizing the need to respect the intellectual property of online content creators. Researchers must therefore assess whether the targeted content is protected by copyright and, if so, whether it falls within any applicable exceptions, such as the use for scientific research.

Third, contract law is a critical consideration, as the terms and conditions (ToS) of social media platforms generally prohibit automated data collection without explicit permission. For example, platforms like TikTok, Facebook, Instagram, and YouTube explicitly restrict data scraping in their ToS, as outlined in the previous table. Violating these terms can lead to account bans, legal claims, and potential liabilities under both GDPR (e.g., unauthorized processing of personal data under Articles 6 and 9) and contract law.

If, after careful consideration, the decision is made to proceed with data scraping, researchers must ensure that they have the appropriate legal permissions and ethical approvals in place, in parituclar CUNI who is expected to take part into these data collection activites. This includes applying for API access where available and securing ethical clearance from their own institutional ethics committees. Each partner organisation is responsible for ensuring that the proposed data collection complies with its internal guidelines, legal obligations, and the broader ethical commitments of the TWIN4DEM project. This approach helps ensure consistency, accountability, and alignment with both institutional and project-level standards.

5. Authorship Guidelines

To support TWIN4DEM's aim to transparency, fairness, and consistency across all research outputs produced within the TWIN4DEM consortium, this section aims to reflect on the shared commitment to ethical research, acknowledges the value of diverse contributions, and promotes practices that are both inclusive and career supportive. While disciplinary conventions may vary, the goal is to create a shared understanding of what constitutes meaningful contribution and how that contribution should be recognized.

TWIN4DEM brings together partners from a wide range of disciplines and methodological backgrounds. Given the interdisciplinary nature of the project, and the variety of outputs the project is expected to produce, it is essential that authorship discussions happen early and openly. These guidelines aim to encourage thoughtful and constructive conversations about authorship decisions. They also aim to harmonize best practices drawn from editorial standards such as Springer's authorship principles, Elsevier's CRediT taxonomy for contributor roles, targeted journals such as West European Politics, Perspective in Politics, Journal of Artificial Societies and Social Simulation, and Guidelines on the Commission of Publication Ethics. As such, this section aims to provide several scenarios that harmonize authorship practices within the consortium.

5.1 Guiding principles

The authorship approach adopted by the TWIN4DEM consortium is grounded in principles of clarity, fairness, and mutual respect among researchers contributing to research outputs. First, it is essential to recognise individual contributions to published work. This goes beyond assigning credit, it is also a way of valuing the many forms of effort that contribute to collaborative research, from conceptual design and data analysis to writing, coordination, and review.

Given the interdisciplinary nature of the project and the variety of institutional contexts involved, the consortium recognises the importance of reducing potential authorship disputes. Early, transparent discussions about roles, expectations, and authorship decisions are strongly encouraged throughout the development of any joint output. These conversations are key to building a collaborative environment where contributors feel respected and acknowledged.

To support transparency, the consortium encourages the use of contribution statements, particularly those based on the CRediT taxonomy. This approach follows principles established by several journals and offers a clear way to specify who is in charge of each part of the work. These statements should as much as possible be included during the submission process and appear in the final publication.

Responsibility and accountability are also central to ethical authorship. As highlighted in several editorial policies, all authors are expected to review and approve the final version of the manuscript and to accept shared responsibility for its content. Authorship implies recognition and commitment to upholding the integrity of the research.

While these internal guidelines are meant to harmonise practices within the TWIN4DEM consortium, they are not meant to override the policies of the journals and conferences targeted for publication. If a journal or conference provides specific authorship rules regarding order, acknowledgments, or contribution formats, those will take precedence.

5.2 Authorship criteria

What constitutes an author? In the context of TWIN4DEM, authorship must reflect a meaningful and direct contribution to the research output. A person should be considered an author if they have made a substantial contribution to the design of the study, the acquisition or analysis of data, or the development of research content. In addition, authorship also requires active involvement in drafting or critically revising the substantive content (including written text, code, data models, algorithms, or other significant technical components) of the publication and formal approval of the final version prior to submission. These elements, taken together, constitute the basis for being recognised as an author. Individuals who support the project in other valuable ways, such as through administrative assistance, provision of funding, or general supervision, must be acknowledged in a dedicated section.

An important distinction in authorship is the role of the corresponding author. This person is in charge of taking a leading responsibility for communication, submission, peer review and publication process. This role serves a functional responsibility. For practical reasons, it is advised to designate someone whose contact details are unlikely to change over (and possibly after) the publication process and who commits to stay available over the course of the peer-review process. As a best practice, the order of authors should be determined collectively among contributors and agreed upon in advance of submission. Where necessary, this can be revisited during the drafting process if the scope of contributions changes or unexpected factors come into play.

5.3 Order of authorship

Regarding authorship order, the first author position is generally considered the most visible and is often associated with the researcher who leads the work, unless authors are ranked alphabetically (see below). The final author position can, in some disciplines, indicate senior oversight or leadership. The order in which authors appear on a publication should reflect a shared understanding among contributors of their

relative contributions. As noted in the ICMJE guidelines, authorship order should be a joint decision, and authors must be prepared to explain how that order was determined. However, the guidelines leave flexibility for research groups to agree on their own arrangements.

In some disciplines, alphabetical ordering is common practice, especially when contributions are equal. In such cases, it is advisable to use the Credit taxonomy to detail the contribution of each author. Other disciplines may follow a more hierarchical structure where the first author is understood to have led the work and the last author often reflects supervisory or senior roles. The consortium recognises the importance of visibility for early-career researchers and encourages, when appropriate, the prioritisation of junior colleagues for lead authorship in cases of comparable contribution, only when all of the parties agree. These decisions should be made openly and in agreement with all involved contributors.

5.1.1 Scenario 1 – Where the order reflects the level of contribution

In scenarios where the order of authorship is closely tied to the nature of each contribution, a formal author contribution statement is strongly recommended. This can take the form of a short paragraph that outlines who was involved in which part of the work or be presented using the CRediT taxonomy, which is now widely accepted across many journals. Using structured templates or tables helps make individual contributions visible and can be particularly useful in demonstrating fairness in collaborative outputs.

Table 2 CRediT Taxonomy table, to be filled out by partners

Term	Definition	Responsible Party
Conceptualization	Ideas; formulation or evolution of overarching research goals and aims	*Name, Institution*
Methodology	Development or design of methodology; creation of models	*Name, Institution*
Software	Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components	*Name, Institution*

Validation	Verification, whether as a part of the activity or separate, of the overall replication/ reproducibility of results/experiments and other research outputs	*Name, Institution*
Formal analysis	Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data.	*Name, Institution*
Investigation	Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection	*Name, Institution*
Resources	Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools	*Name, Institution*
Data Curation	Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later reuse	*Name, Institution*
Writing - Original Draft,	Preparation, creation, and/or presentation of the published work, specifically writing the initial draft (including substantive translation)	*Name, Institution*
Writing - Review & Editing,	Preparation, creation, and/or presentation of	*Name, Institution*

	the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post publication stages	
Visualization	Preparation, creation and/or presentation of the published work, specifically visualization/ data presentation	*Name, Institution*
Supervision	Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team	*Name, Institution*
Project administration	Management and coordination responsibility for the research activity planning and execution	*Name, Institution*

Note: Only 13 roles instead of 14 were included due to funding already being established.

To support transparency and consistency across the consortium, author contribution statements should be based on CRediT. This taxonomy provides a clear structure to describe each person's role in the research process, covering a range of contributions such as conceptualization, methodology, software development, data curation, analysis, writing, and supervision. The statement can take the form of a structured list or a narrative paragraph, depending on the journals or conference's requirements. Below are two examples of how this might look in practice:

a. A.B., C.D., and E.F. designed the framework for the project A.B., G.H., and I.J. led the collection and harmonisation of survey and institutional data. M.N. and P.Q. conducted the focus groups and contributed to stakeholder mapping. C.D.,

E.F., and R.S. developed the policy scenarios. A.B., C.D., and E.F. drafted the manuscript, with all authors reviewing and contributing feedback.

b. A.B: Conceptualization, Methodology, C.D.: Data Curation, Formal Analysis. E.F: Software and Modelling. G.H. Supervision.: Writing- I.J. Reviewing and Editing,

5.1.2 Scenario 2 – Order does not necessarily reflect the level of contribution

When contributions are equal or similar, and there is agreement among authors, the project may support prioritising early-career researchers for first authorship where this may contribute to their professional progression. Such decisions should be made on a case-by-case basis, guided by openness and mutual respect among contributors. Regardless of the final arrangement, clarity with editors and within the consortium is essential.

5.3 Disclosure

All authors are expected to include a disclosure statement outlining any sources of funding in accordance with Article 17 of the Grant Agreement. In conference papers, this statement often goes in the “Acknowledgment” section. This is a standard part of responsible publishing and contributes to the transparency and credibility of the research and aims to avoid conflict of interest. For outputs related to TWIN4DEM, an appropriate disclosure statement might read:

> *Funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.*

5.3 Disputes

In the case of an authorship dispute, it is expected that the issue will first be addressed in good faith among the contributors involved. If the matter cannot be resolved informally, it should be brought to the attention of the project coordinator, who will facilitate a solution in consultation with the relevant work package leads and, if necessary, the institutional policies of the organisations concerned. By contributing to TWIN4DEM publications, all authors agree to uphold the principles outlined in these guidelines and to engage in authorship decisions that support transparent, equitable, and ethically grounded research practices.

6. Ethics in dissemination strategies

Dissemination is a core component of the TWIN4DEM project. It ensures that research results, tools, and lessons learned reach the broader community, including researchers, developers, policymakers, and the public. However, dissemination also carries ethical responsibilities, especially when it involves sharing information derived from human participants, AI systems trained on sensitive data, or materials produced through real-world testing.

When using images, audio, or video recordings of individuals, explicit and informed consent must be obtained in advance. This includes presentations at conferences, publication on websites, and social media posts. Even when individuals appear in group settings or are not named, ethical use still requires considering whether they could be recognised or affected by the context in which the material is shared. According to Article 5(1)(c) of the GDPR, the principle of data minimisation states that only data which are relevant and necessary for the intended purpose should be collected and used. Moreover, Recital 26 and Article 4(1) of the GDPR stress that if someone can be identified, even indirectly, through available information, the data are still considered personal. This highlights the importance of being mindful of re-identification risks, especially when sharing audiovisual content. Dissemination strategies should therefore include practical steps to reduce these risks and protect individuals' privacy.

When organising events, partners are encouraged to design spaces and documentation practices with ethics in mind. This may include using signage to indicate where photography or filming will take place, offering badges or stickers to identify individuals who do not wish to be recorded, and reminding staff and external service providers of these rules. Images should focus on activities, environments, or materials, rather than faces, unless individuals have clearly agreed to appear and have understood how their image will be used.

TWIN4DEM partners must also avoid disseminating content that could misrepresent findings, overstate results, or obscure risks. Ethical communication means being transparent not only about successes but also about limitations and uncertainties. This is especially important when describing AI functionalities or system capabilities. Where prototypes are still under development, or pilot results are preliminary, this should be clearly stated to avoid creating unrealistic expectations. Finally, before publishing or publicly sharing project outputs, partners must follow internal procedures for approval, and ensure that no sensitive deliverables are disclosed in violation of the classification outlined in the Grant Agreement.

7. Conclusion

The project TWIN4DEM operates within an ethical and legal framework that aligns with standards on human rights, data protection, and responsible research. Safeguards are in place to protect participants involved in focus groups and interviews, as well as the individuals whose data will be used. Operationalising these obligations demonstrates a proactive commitment to transparency, inclusion, and accountability, which is implemented through informed consent procedures and techniques such as data anonymisation and minimisation.

Ethical support is also provided throughout the entire project, including participant recruitment and the recognition of researchers' authorship in the dissemination of findings. It is essential that all partners are actively involved in the ethical development of the project, as this is the most effective way to ensure compliance and to foster awareness around responsible research practices.

8. References

Albert, T., & Wager, E. (2024, December 10). *How to handle authorship disputes: a guide for new researchers*. COPE: Committee on Publication Ethics. https://publicationethics.org/guidance/guideline/how-handle-authorship-disputes-guide-new-researchers

Cambridge Core. (2023). *Preparing your materials*. Cambridge Core. https://www.cambridge.org/core/journals/perspectives-on-politics/information/author-instructions/preparing-your-materials

Council of Europe. (n.d.). *Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data*. Impact of the European Convention on Human Rights. https://www.coe.int/en/web/impact-convention-human-rights/convention-for-the-protection-of-individuals-with-regard-to-automatic-processing-of-personal-data#/

Department of Health, Education, and Welfare. (1979). *The Belmont Report*. https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf

European Commission. (2018a, August 22). *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01)*. Ec.europa.eu. https://ec.europa.eu/newsroom/article29/items/612053/en

European Commission. (2018b, August 22). *Guidelines on Transparency under Regulation 2016/679* . Ec.europa.eu. https://ec.europa.eu/newsroom/article29/items/622227

European Commission. (2020, July 17). *Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment | Shaping Europe's digital future*. Digital-Strategy.ec.europa.eu. https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

European Data Protection Board. (2020, May 4). *Guidelines 05/2020 on consent under Regulation 2016/679 | European Data Protection Board*. Www.edpb.europa.eu. https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-052020-consent-under-regulation-2016679_en

European Parliament and of the Council. (2016, April 27). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Europa.eu. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679

European Parliament and of the Council . (2016, April 27). *Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision*

2008/977/JHA. Europa.eu. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016L0680

JASSS. (2025). *JASSS: How to submit a paper*. Jasss.org. https://www.jasss.org/admin/submit.html

Taylor and Francis Group. (2025, May 12). *Defining authorship in your research paper Co-authors, corresponding authors, and affiliations*. Author Services; AS. https://authorservices.taylorandfrancis.com/editorial-policies/defining-authorship-research-paper/?_gl=1

The European Commission. (2025, May 6). *expert group on AI | Shaping Europe's digital future*. Digital-Strategy.ec.europa.eu. https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai

World Medical Association. (1964). *Declaration of Helsinki*. World Medical Association. https://www.wma.net/what-we-do/medical-ethics/declaration-of-helsinki/doh-jun1964/

9. Annex 1: Draft template of information sheet and consent form of focus groups

Participant information sheet about TWIN4DEM and informed consent form

Project Overview: reasons for collecting personal data and funding source

Democracy International as part of the consortium of eleven partners, is undertaking a TWIN4DEM project to prototype digital twins of four European democratic systems (Czechia, France, Hungary and the Netherlands). We are asking you to help with this TWIN4DEM activities by participating in the focus group activities and letting us analyse information about your input and socio-demographic information collected above. [Partners will briefly state why the validation is important in the context of the project]

TWIN4DEM is funded by the European Commission's Horizon Europe program, through Grant Agreement no. 101178061. You can find more information about TWIN4DEM project here: https://twin4dem.eu/

Reasons why you are being invited to participate

You are being invited to participate in this research because your expertise and professional experience are directly relevant to the project's focus on democratic processes and democratic resilience. Participants are selected based on their knowledge and engagement in these areas, with the goal of ensuring fair and consistent selection across all countries involved. The focus groups aim to reflect a diverse range of perspectives, including different backgrounds, genders, ages, ethnicities, political beliefs, and socioeconomic contexts. Policymakers and stakeholders from both ruling and opposition parties may be included, while maintaining a balanced and neutral environment for open dialogue.

Benefits of participation

You will receive150 euros for your time as well as a travel reimbursement.

Data Controller and Data Protection Officer (DPO)

The Data Controller responsible for processing your data is: [Partners will add the relevant DPO depending on the country of focus groups]

- Name:
- Data Protection Officer name:
- DPO contact information:

Purpose of data processing

Your data will be collected to [explain purpose and procedures. E.g. your data will be use to analyse x and y. This process will be done through a focus group that will x and y]

Data being collected

The data we will collect is [x and y, maybe voice recording, photos, or something else]. Researchers may also collect other types of personal data for administrative purposes related to the research, such as this informed consent form. These categories of personal data include: name and surname, occupation, gender, age group and ideological or political background. All data collection will follow minimization procedures, meaning, we will only recollect what is strictly necessary.

Information that you provide to us will be recorded via video and stored securely at the Democracy International's storage system. We understand that some information may be sensitive and we will keep your information confidential and use it only for TWIN4DEM.

Legal basis for personal data processing

The legal basis for processing your personal data is your explicit consent, through this document, pursuant to Articles 6(1)(a) and 9(2)(a) of the European Union's General Data Protection Regulation (https://eur-lex.europa.eu/eli/reg/2016/679/oj).

Information about data sharing

This project is carried out within a consortium of partners. This means that it may be necessary for the Data Controller to share the information you provide with other TWIN4DEM project partners. Any sharing of your personal data will be strictly limited to partners who need access for research purposes, and your data will not be shared outside the consortium. It will not be published or disclosed in any way that could identify you.

Data retention period

Your personal data will be stored on the servers and facilities of the Democracy International for [X] years, then destroyed. The TWIN4DEM consortium may make some of the anonymized data collected during this project publicly available for educational or research purposes, including in areas beyond the scope of this study. Any data shared will be carefully anonymized to ensure that it cannot be traced back to you, and no personal information or identifying details will be disclosed.

Data subject rights: Under GDPR, you have the right to:

- Obtain information about whether, how, and why we process your personal data
- Access your personal data
- Rectify or correct your personal data
- Have your personal data erased
- Restrict further processing of your personal data
- Request the transfer of your personal data in an electronic and structured form to you or to another party (right to “data portability”)
- Lodge a complaint with a supervisory authority
- Withdraw your consent, at any time, by sending an e-mail to [DPO’s email]

Informed Consent

Please read each statement carefully and initial only the box for the statement you consent to.

No.	Statement	Initials
1	I confirm that I have read and understood the information provided above. I have had the opportunity to consider the information, ask questions, and receive satisfactory answers.	

2	I understand why I was selected, that my participation is voluntary and that I am free to withdraw at any time, without any adverse consequences by contacting the Data Protection Officer.	
3	I understand that my data may be processed by the several partners of the TWIN4DEM consortium.	
4	I understand that my data will be used for research purposes, and that other personal data will be processed for administrative purposes related to this research.	
5	I consent to having my inputs recorded and images taken for dissemination purposes.	
6	I understand that I can exercise my data protection rights at any time, including access to, correction of, erasure of, transfer of, and the restriction on further processing of my personal data.	
7	I understand that my data will be stored by DI for [X] years. After this period, it will be destroyed.	

Name of Participant	Date (dd/mm/yy)	Signature
_____	_____	_____

Name of person taking consent		
_____	_____	_____