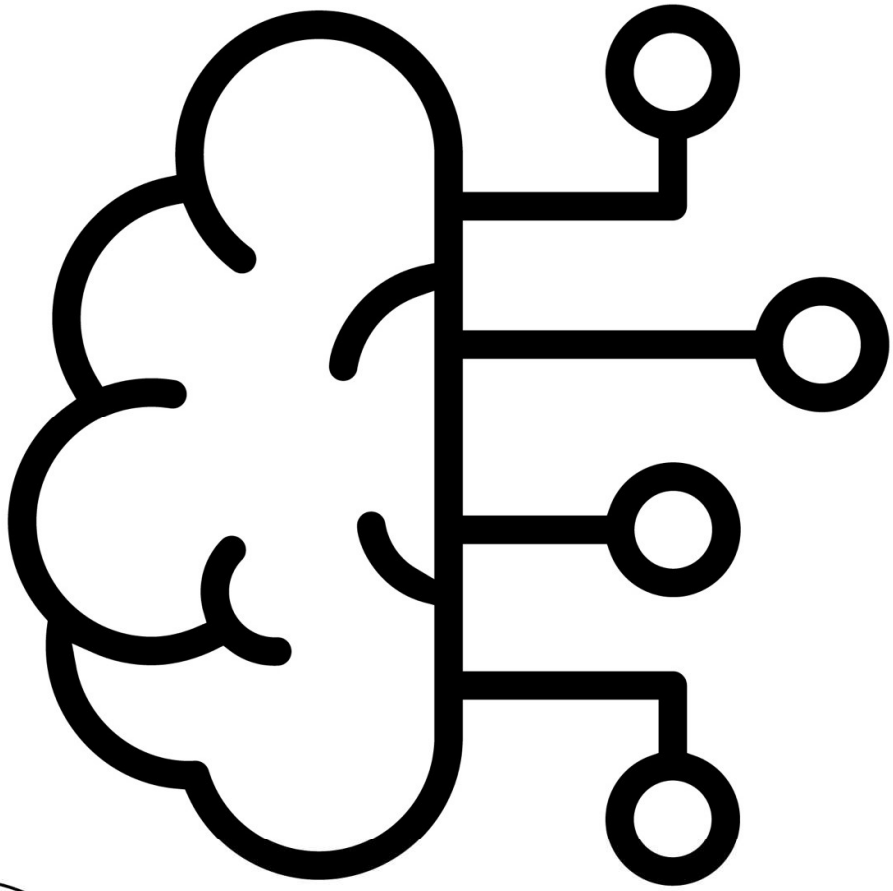
MRS Evidence Matters™

MRS Guidance on Using AI and Related Technologies

April 2025



MRS Campaign for Better Data

Introduction

MRS has produced this guidance to help practitioners act legally and ethically when using AI and related technologies.

This is the second iteration of this document, and the intention is that the guidance will be updated and enhanced regularly to keep abreast of this rapidly evolving policy area, particularly as regulatory frameworks develop.

MRS has referred to other AI guidance documents and ethical frameworks in compiling this guidance to ensure that the MRS approach is compatible with other business, legal and societal approaches.

Scope

This guidance apples to all MRS members and MRS Company Partners and should be read in conjunction with the MRS Code of Conduct.

Interpretation of Requirements

When requirements use the word "must" these are mandatory requirements and is a principle or practice that applies the MRS Code of Conduct, which Members and Company Partners are obliged to follow.

The requirements which use the phrase "should" describe implementation and denotes a recommended practice. "May" or "can" refer to the ability to do something, the possibility of something, as well as granting permission.

Explanation of Key Terms from the MRS Code of Conduct

The relevant definitions from the MRS Code of Conduct (2023) are:

Child: A child is an individual under the age of 16.

Client: includes any individual, organisation, department or division, including any belonging to the same organisation as an MRS Member, which is responsible for commissioning or applying the results from a project.

Data: is information collected in any nature or format.

Data analytics: is the process of interrogating data to identify patterns, correlations, trends or other information. This also includes modelling, forecasting and aggregation of data.

Participant: is any individual or organisation from or about whom data is collected.

Practitioners: includes all individuals within the research, insight and data analytics supply-chain e.g., researchers, software suppliers, project managers, moderators, contractors, freelancers and temporary workers.

Reasonable action: is such actions as a person in their position (in light of experience, role, responsibilities etc.) should be expected to take to adhere with the provisions of the MRS Code.

Research: is the collection, use, or analysis of information about individuals or organisations intended to establish facts, acquire knowledge or reach conclusions. It uses techniques of the applied social, behavioural and data sciences, statistical principles and theory, to generate insights and support decision-making by providers of goods and services, governments, non-profit organisations and the general public.

Special category data: is the processing reveals racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union Membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation.

Stakeholders: individual or group that has an interest in any decision or activity being undertaken.

Vulnerable people: Vulnerable people means individuals whose permanent or temporary personal circumstances and/or characteristics mean that they are less able to protect or represent their interests (see MRS Best Practice Guide on Research Participant Vulnerability).

AI Definitions and Related Terms

AI exists on a spectrum which ranges from automation to innovation. In the context of research and insight the following terms and definitions are the most applicable:

Algorithm: a process or set of rules to be followed in calculations or other problem-solving operations. [ISO 20252: 2019 definition]

Artificial Intelligence (AI): a system where machine or software systems make autonomous decisions normally requiring human intelligence and acts, creates, evolves, or changes.

Generative AI: a type of AI that is able to produce new content (text, images, audio, synthetic data or other media), based upon input training data, in response to prompts. LLMs and Generative AI are two related but

distinct areas of AI. Generative AI is a broader concept referring to any machine learning model capable of creating output after training.

Large Language Models (LLMs): a machine learning model which processes huge amounts of existing information to understand, summarise, generate and predict new content. Examples of LLMs include Chat GPT, DeepSeek-R1, Gemini and Claude.

Machine learning (ML): the use and development of digital systems that use algorithms and statistics models to analyse and draw inferences from patterns of data. As new data is gathered the models can be adapted to enable algorithms to learn using data to improve future performance on specified tasks.

Natural Language Processors (NLPs): a field of AI focusing on human languages. LLMs are specific models used within NLP that specialise in language related tasks.

Synthetic data: data that is artificially created to augment or replace real data. Synthetic data can be used to create models that are trained to reproduce the characteristics and structure of the original data producing similar results to the original data.

For the purpose of this guidance when using the term "AI and related technologies" assume the guidance applies to any of the above technologies.

AI and Research, Insight and Data Analytics

Within the research, insight and data analytics sector, the types of activities that could be undertaken, to some degree via the use of current AI and related technologies includes operations, data generation, data collection, data measurement and analysis, reporting and report writing.

The AI series of MRS Delphi Reports and the reports from the MRS Advanced Insight and Analytics (AIA) Council provides more detailed case studies of how AI is being used in research and insight.

Legal and Regulatory Obligations

UK AI legislation

The UK has so far implemented a pro-innovation approach to AI, with a network of existing regulators overseeing the use of AI within their sector-specific domains e.g., the ICO focusing on AI's data protection and privacy considerations. Against this backdrop the UK Government is drafting an AI bill focusing on 'frontier AI', strengthening voluntary AI safety requirements to become legally binding and to strengthen the current UK Government's AI governance arrangements.

MRS will provide an interpretation of the developments and requirements as they evolve. In the meantime, MRS is making direct representations to the UK Government on this issue. The MRS responses to the UK Government's White paper and other related consultations are available here.

AI and UK Data Protection Legislation

The Information Commissioner's Office (ICO) has set out how to apply the principles of the UK GDPR to the use of information in AI systems.[1] Data protection legislation applies to all personal information used in AI and related technologies. The ICO has identified the following eight questions when using AI and related technologies:

1. **What is your lawful** basis **for processing personal data?** If you are processing personal data you must identify an appropriate lawful basis, such as consent or legitimate interests.

2. **Are you a controller, joint controller or a processor?** If you are developing generative AI using personal data, you have obligations as the data controller. If you are using or adapting models developed by others, you may be a controller, joint controller or a processor.

3. **Have you prepared a Data Protection Impact Assessment (DPIA)?** You must assess and mitigate any data protection risks via the DPIA process before you start processing personal data. Your DPIA should be kept up to date as the processing and its impacts evolve.

4. **How will you ensure transparency?** You must make information about the processing publicly accessible unless an exemption applies.

[1] See Generative AI: eight questions that developers and users need to ask: https://ico.org.uk/about-the-ico/media-centre/blog-generative-ai-eight-questions-that-developers-and-users-need-to-ask/

If it does not take disproportionate effort, you must communicate this information directly to the individuals the data relates to.

5. **How will you mitigate security risks?** In addition to personal data leakage risks, you should consider and mitigate risks of model inversion and membership inference, data poisoning and other forms of adversarial attacks.

6. **How will you limit unnecessary processing?** You must collect only the data that is adequate to fulfil your stated purpose. The data should be relevant and limited to what is necessary.

7. **How will you comply with individual rights requests?** You must be able to respond to people's requests for access, rectification, erasure or other information rights.

8. **Will you use generative AI to make solely automated decisions?** If so – and these have legal or similarly significant effects (e.g. major healthcare diagnoses) – individuals have further rights under Article 22 of UK GDPR.

More information and guidance from the ICO on AI and data protection can be found here: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/

EU Legislation

The EU's AI Act was published on the Official Journal of the EU on 12th July 2024, coming into force from 1st August 2024. The legislation will be fully applicable 24 months after coming into force, with a graduated introduction as follows:

- **2nd February 2025**: Prohibitions on unacceptable risk AI became applicable.
- **2nd August 2025**: Obligations for general-purpose AI (GPAI) governance become applicable.
- **2nd August 2026**: All rules of the AI Act become applicable (except Article 6) including obligations for standalone high-risk systems (e.g., credit scoring, HR, insurance risk assessment).
- **2nd August 2027**: Obligations for all other high-risk systems under specific sectoral legislation (e.g., medical devices, toys) become applicable (Article 6).

The EU AI Act has global reach beyond the EU and can apply to businesses located outside of the EU, for example businesses which market or utilise AI in the EU or where the output is used in the EU. The key question is the effect of the AI on the EU, and not necessarily where the relevant provider or operator is situated.

The Act states that AI systems that can be used in different applications are analysed and classified according to the risk they pose to users[2]. The different risk levels will mean more or less regulation. The new rules establish obligations for providers and users depending on the level of risk from AI. While many AI systems pose minimal risk, they need to be assessed.

AI systems deemed too risky and a potential threat to people are banned. They include eight types of prohibited practices:

- AI systems for social scoring.
- AI systems to infer emotions in the areas of workplace and education institutions.
- AI systems that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage.
- AI systems linked to assessing or predicting criminal offences, based solely on profiling or assessing personality traits and characteristics.
- Biometric categorisation systems that categorise individuals based on their biometric data to deduce or infer race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation.
- The deployment of subliminal, manipulative or deceptive techniques.
- The exploitation of vulnerabilities due to age, disability or a specific social or economic situation.
- The use of real-time remote biometric identification systems in publicly accessible spaces for law enforcement purposes.

AI systems that negatively affect safety or fundamental rights are considered high risk and are divided into two categories:

1. AI systems that are used in products falling under the EU's product safety legislation. This includes toys, aviation, cars, medical devices and lifts.

2. AI systems falling into eight specific areas that will have to be registered in an EU database:

   - Biometric identification and categorisation of natural persons
   - Management and operation of critical infrastructure
   - Education and vocational training
   - Employment, worker management and access to self-employment

[2] For the full article from which this is based see European Parliament News: https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

- Access to and enjoyment of essential private services and public services and benefits
- Law enforcement
- Migration, asylum and border control management
- Assistance in legal interpretation and application of the law.

All high-risk AI systems will be assessed before being available to the market and also throughout their lifecycle.

Generative AI, like Chat GPT, would have to comply with transparency requirements:

- Disclosing that the content was generated by AI
- Designing the model to prevent it from generating illegal content
- Publishing summaries of copyrighted data used for training

MRS is currently contributing to the development of a Code of Practice for general purpose AI which is being drawn up by the EU's AI office. MRS is inputting into this Code, via its participation in the European Research Federation EFAMRO.

The EU's General-Purpose AI Code of Practice is a requirement of the AI Act and will become effective from August 2025. The Code is due to be finished by May 2025. This MRS AI guidance takes a complementary approach to the EU and the two documents should be compatible.

AI Ethics and the MRS Code of Conduct

AI ethics are the parameters and guardrails that determine the design, use and outputs from the use of AI and go beyond the legal requirements. Ethics are essential when using AI and related technologies since the foundation for AI and related technologies is data, much of which is drawn from human behaviour. One of the key weaknesses of AI and related technologies is the reliance on input data of earlier large data systems, particularly owing to the level of inaccuracy and frequent bias of the input data. Research and insight produced by AI and related technologies can be a start-point, which requires human expertise and practitioner input to check and validate the content.

If AI and related technologies are used without consideration for ethics, the output from AI can amplify and emphasize human biases and inaccuracies which could result in harm to individuals, business, and society. Conversely, if AI is positively managed with ethical principles at its core the potential of its possibilities can be maximised.

Ethics are essential for any profession, and any approach by MRS needs to complement the MRS Code of Conduct. Within the MRS Code of Conduct there are 12 ethical principles which underpin all rules and requirements within the Code. These principles apply to the use of AI in all its forms. The following details the MRS Code of Conduct principles, setting out how these apply to the related ethical principles when using AI and related technologies. The intention is that the MRS ethics guidance will be updated regularly to reflect use cases and understanding.

The MRS AI ethics approach is structured around four ethical pillars across the 12 principles of the MRS Code:

Communication, Use and Access	**Client Data and Confidentiality**	**Data Protection and Privacy**	**Reputation of the Profession and the Sector**
– Transparency – Explainability – Accessibility and understandability – Fairness and Impartiality	– Responsibility and Ownership – Appropriateness – Human Oversight	– Privacy and Security – Dignity and Autonomy – Proportionality and Robustness	– Awareness and Literacy – Trust – Sustainability

Communication, Use and Access

MRS Code Principle 1: MRS Members shall ensure that their professional activities can be understood in a transparent manner.

MRS Code Principle 2: MRS Members shall be straightforward and honest in all professional and business relationships.

MRS Code Principle 3: MRS Members shall be transparent as to the subject and purpose of data collection.

MRS Code Principle 4: MRS Members shall ensure that their professional activities are not used to unfairly influence views and opinions of participants.

When using AI and related technologies, applying the above four principles from the MRS Code of Conduct requires practitioners to adhere to the following ethical requirements:

- **Transparency:** To reinforce trust, users (participants, clients, colleagues) must be able to determine how a service, system or tool works; be able to evaluate its functionality, how and why it is being deployed and comprehend its strengths and limitations. The communication must be appropriate and suitable for the audience e.g., more detailed for research commissioners, sufficient to enable informed consent for participants, etc. Similarly, when AI and related technologies are being used within professional activities this must be transparent to users of any resulting AI generated outputs including clients.
- **Explainability:** Any AI tool, system or use must be explainable in terms of how, for example, algorithms and systems are used, assumptions have been made and applied and resulting adaptions.
- **Accessibility and Understandability:** When providing information to explain and make transparent the use of AI and related technologies the information being supplied must be accessible and understandable across a range of stakeholders with a range of knowledge including participants, clients and colleagues.
- **Fairness and Impartiality:** When designing, using and/or calibrating AI and related technologies systems must be structured to enable humans to make fair choices and to not influence opinions or activities. The systems must also be fair and inclusive to all participants, without bias favouring or disadvantaging particular groups of participants or individuals. This should be considered against internal and external initiative such as the MRS Inclusion Pledge and the MRS Representation in Research initiatives.

Client Data and Confidentiality

MRS Code of Conduct principle 5: MRS Members shall respect the confidentiality of information collected in their professional activities.

MRS Code of Conduct principle 9: MRS Members shall exercise independent professional judgement in the design, conduct and reporting of their professional activities.

When using AI and related technologies, applying the above two principles from the MRS Code of Conduct requires practitioners to adhere to the following ethical requirements:

- **Responsibility and Ownership:** When using AI and related technologies it must be clear who is responsible for the design, development and deployment of the technology. AI systems must be auditable and traceable with ownership of any data outputs identified and attributed. Practitioners must protect all data used in research, insight and analytics (e.g., client and participant data) and are responsible for protecting the confidentiality of such data.
- **Appropriateness:** AI and related technologies must be appropriate for the purpose. When using AI and related technologies, different purposes have different risks, different AI and related technologies also have different levels of risk and these must be reflected in the approach taken.
- **Human oversight:** When using AI and related technologies there must be some element/s of human oversight within the process to ensure greater accuracy and safety, and a robust approach to AI governance, decision making and maintenance. There must be oversight, impact assessments, testing and replicability processes, and audit and due diligence mechanisms in place to avoid conflicts with human norms, client requirements and participant expectations.

Privacy and Security

MRS Code of Conduct principle 6: MRS Members shall respect the rights and well-being of all individuals.

MRS Code of Conduct principle 7: MRS Members shall ensure that individuals are not harmed or adversely affected by their professional activities.

MRS Code of Conduct principle 8: Members shall balance the needs of individuals, clients, and their professional activities.

When using AI and related technologies, applying the above three principles from the MRS Code of Conduct requires practitioners to adhere to the following ethical requirements:

- **Privacy and Security:** AI and related technologies must prioritize and safeguard participants' privacy and data rights and provide explicit assurances to users about how personal data will be used and protected including robust security measures.
- **Dignity and Autonomy:** Practitioners and users of AI and related technologies must respect and protect the fundamental freedoms and human dignity of all participants, including basic human rights and autonomy.
- **Proportionality:** The use of AI and related technologies must be proportionate and not go beyond what is necessary to achieve any legitimate business aim or objective. AI and privacy risk assessments must be undertaken to identify and mitigate risk and prevent harms which may result.
- **Robustness:** AI and related technologies must be robust to protect from data attacks, minimising security risks and enabling confidence in outputs and outcomes.

Reputation of the Profession and the Sector

MRS Code of Conduct principle 10: Members shall ensure that their professional activities are conducted by persons with appropriate training, qualifications and experience.

MRS Code of Conduct principle 11: Members shall protect the reputation and integrity of the profession.

MRS Code of Conduct principle 12: Members shall take responsibility for promoting and reinforcing the principles and rules of the MRS Code of Conduct.

When using AI and related technologies, applying the above three principles from the MRS Code of Conduct this requires practitioners to adhere to the following ethical requirements:

- **Awareness and Literacy:** Practitioners using AI and related technologies must have appropriate training and understanding of the systems and processes they are using. This should include AI ethics training.
- **Trust:** Practitioners rely upon the trust of participants and clients to undertake their professional activities. When using AI and related technologies practitioners must follow the MRS Code and this guidance to ensure that trust is maintained in the profession.
- **Sustainability:** The sustainability and impact of AI and related technologies should be considered and understood against any internal and external goals (such as the MRS Climate Pledge).

Requirements

The following are the MRS rules and requirements when using AI and related technologies.

General Rules of Professional Conduct

1. Practitioners must ensure that their professional activities using any AI and related technologies conforms to the national and international legislation relevant to any given project, including UK, EU and any other international data protection and AI legislation.

   Comment: When using third party solutions and systems which incorporate AI and related technologies practitioners should obtain information about the solutions to determine their suitability for use.

2. Practitioners must ensure that they adhere to all relevant legal and ethical requirements when conducting their professional activities including the use of AI and related technologies.

   Comment: Practitioners should create internal AI and related technologies usage policies and training to instruct staff, sub-contractors and any other stakeholders about the acceptable and unacceptable use of AI and related technologies. These policies should be monitored to ensure that individuals do not breach or cause a breach of these guidelines and/or the MRS Code of Conduct.

3. Practitioners must determine when the use of AI and related technologies are appropriate for any given project.

   Comment: Whilst the use of AI and related technologies are the dominant conversation within research, these technologies may not always be appropriate, and it is for practitioners to use their judgement to determine when and where it is most suitable.

4. Practitioner must take reasonable steps to ensure that the use of AI and related technologies are proportionate and do not go beyond what is necessary to achieve any agreed business aims or objectives.

5. Practitioners must inform clients, participants, staff, sub-contractors and any other stakeholders when AI and related technologies are being used in any part of a project where the use of AI and related technologies are either classified as 'high risk' and/or affect research outputs, participants, clients, and/or decision making.

   Comment: This information can be relayed in standard terms and conditions (for example with sub-contractors), particularly if the use of AI and related technologies is not particularly intrusive, does not include personal data and/or client data and is not likely to have a significant impact on the research outputs. However, where AI and

related technologies are core to a process this needs to be communicated more clearly e.g., within proposals, client briefs.

6. Practitioners must inform clients of the limitations of any AI and related technologies which are being used for their projects, including any risks.

   Comment: Communications about limitations could include issues such as concerns about data provenance and ownership, the lack of transparency of sources, level of reliability or completeness of data, the potential for bias, etc.

7. Practitioners must ensure that communications about AI and related technologies include information about who is responsible for the design, development, ownership and deployment of the technology and tools.

   Comment: It is essential that there is a clear understanding of ownership and responsibility for any AI and related technologies, particularly if either participants or clients request such information. It is also necessary to have the information to fulfil any requests such as data deletion or data rectification requests from participants. Any assurances provided to participations, e.g., in response to data deletion or data rectification requests, need to be honoured. As such it is recommended that when relying upon AI and related technology owners and providers to fulfil such requests, details of the process are obtained before assurances are given (e.g., will personal data be permanently deleted or archived, will data be corrected within the models?).

8. Practitioners must retain adequate records to ensure that AI and related technologies are auditable, replicable and traceable. These records should be used to answer questions about the provenance, accuracy and reliability of records, data or results produced using AI and related technologies.

   Comment: Currently not all AI data (for example from some Generative AI) is traceable to sources. Data from such sources should not be used where there is a doubt about data provenance.

9. Practitioners must ensure that there is human oversight when using AI and related technologies, assessing the quality of research outputs and insights generated. Practitioners must ensure that this human oversight also includes responsibility for safety, robustness, decision making, governance and maintenance of the systems.

   Comment: Activities such as impact assessments, audit, continuous monitoring and improvement and due diligence mechanisms should be considered to meet the above requirements.

10. Practitioners must ensure that those who are using AI and related technologies have appropriate training and understanding of related systems, policies and processes before they are used.

    Comment: It is recommended that AI training includes AI ethics training including covering this guidance.

11. Practitioners should consider and understand the sustainability and impact of AI and related technologies when being used.

    Comment: Many businesses may have sustainability and ESG targets, which could be affected as a result of high usage of AI and related technologies. The sustainability impact should be considered against internal and external goals such as the MRS Climate Pledge.

Commissioning and Design

12. Practitioners must endeavour to provide comprehensive information about how AI and related technologies, systems or tools work including information about functionality, why it is being deployed and an assessment of the advantages and disadvantages of using the tools or technologies.

    Comment: It is appreciated that these requirements may be challenging, particularly if information is not available for AI and related technologies being used. Practitioners are advised to obtain as much information as they can about the technologies they intend to use before deployment.

13. When providing information about AI and related technologies, systems or tools, practitioners must take reasonable actions to ensure that the information is understandable, accessible and explainable in terms of algorithms and systems used, assumptions, adaptions and any other information which is necessary to enable understanding by a wide range of stakeholders.

    Comment: Practitioners should consider different approaches to communicating information about AI and related technologies, similar to the approach used when communicating information about data protection and privacy.

    The key is tailoring such information to the intended audience/s. Practitioners should put themselves into the position of the individuals who will be consuming the information and the level of knowledge of the intended audience. Practitioners should consider splitting communications about AI and related technologies into different categories for different stakeholders and provide tailored AI information for each group. For example, top level communications which summarise overall what AI and related technology is being

used, for what purpose etc. could be suitable for participants whilst more detailed and technical information could be suitable for client compliance teams.

See the ICO and Alan Turing Institute guidance, *Explaining Decisions Made with AI* for further guidance.

14. Practitioners must take reasonable actions to ensure that when designing, using or calibrating AI and related technologies the systems are structured to enable participants to make fair choices and are not designed to influence opinions or activities.

    Comment: See MRS Code of Conduct rule 28 on more detailed guidance about data collection which is relevant for some of this requirement.

15. Practitioners must take reasonable actions to ensure that when designing, using or calibrating AI and related technologies the approach they use is fair and inclusive for all participants.

    Comment: Sectoral inclusivity approaches such as the MRS Inclusion Pledge and the MRS Representation in Research initiatives should be considered in this context.

Confidentiality

16. Practitioners must protect any employer, client and participant supplied data including client and/or employer-owned intellectual property when using AI and related technologies.

    Comment: The use of some AI and related technologies is well-established. However, many of these systems use data input by users to train and refine their systems. If, for example, staff input commercially confidential client data into AI systems, which are not protected with user licences, to brainstorm research design ideas this activity would be in breach of this requirement and in all likelihood would be in breach of client contracts and also data protection legislation depending on the type of data shared. Staff should be made aware of what they can and cannot do with employer, client and participant data (see requirement 2).

17. Practitioners must not share any client data, personal data or confidential information with AI and related technologies which are used to train AI and related technologies systems unless expressly agreed with clients.

    Comment: Practitioners should consider obtaining user licence agreements when using AI and related technologies. This will enable practitioners to set agreed terms of use of such systems, helping to

ensure that staff do not inadvertently mis-use systems and/or any confidential or personal data. In addition it is recommended that guardrails and default settings are checked, even with systems under licence, to ensure that features like 'improve the model for everyone' are disabled.

18. Practitioners must ensure that the rights and responsibilities of themselves, clients and sub-contractors are governed by a written contract, and this includes the use of AI and related technologies including any use of client supplied data.

    Comment: Clients are likely to have Acceptable Use Policies for AI and related technologies and practitioners need to ensure that these align with any intended use of AI for professional purposes.

Children and Vulnerable People

19. Practitioners must ensure that when gathering the permission of a responsible adult, to enable a child to participate in their professional activities, information is relayed about the use of any AI and related technologies within the project.

20. Practitioners must ensure, when gathering consent and before any data is processed using AI and related technologies, that information about the use of such technologies is transparent and accessible by the responsible adults and/or the child (depending on the age of the child).

21. Practitioners must take reasonable steps to assess, identify and consider the particular needs of vulnerable people involved in their professional activities and how their rights are being protected when AI and related technologies are being used.

22. Practitioners must ensure that all participant's rights (e.g., the right to anonymity, the right to withdraw, etc.) are protected and are not weakened as a result of the use of AI and related technologies.

    Comment: See the MRS Code of Conduct for more details about participant rights.

Data Creation and Use

23. Practitioner must obtain participant consent for any personal data used to enhance AI and related technologies, including in the creation of synthetic data models.

    Comment: In addition to consent, participants should anonymise data before it is used in AI and related technologies to protect participants and reduce the likelihood of harm.

24. Practitioners must obtain client consent for the use of AI and related technologies in data creation e.g., the use of synthetic data techniques. When obtaining consent, practitioners must include a clear description of purpose, where in the process AI and related technologies are used, and the abilities and limitations for the use of any proposed AI and related technologies.

25. Practitioners must maintain records detailing how AI and related technologies are used in data creation including how data is generated (e.g., prompts) and combined, how bias and errors are mitigated, the frequency of data updates/refresh and the source of any training data.

    Comment: For example, when synthetic data is used, including when it is combined with non-synthetic data in the same dataset, each synthetic data element would need to be clearly identifiable.

Data Collection

26. Practitioners must inform participants, prior to data collection, when any AI and related technologies are being used for processing personal data and what the consequences are of such use to participants and their personal data.

    Comment: This applies in instances where AI and related technologies are being used in a way which participants might not expect. This does not include instances where low risk everyday AI and related technologies are being used such as when collecting cookies or sending emails, or when AI and related technology systems are being used to automate processes and procedures when no personal data (e.g., from participants) is being processed.

27. Practitioners must ensure that all of their professional activities, including when using AI and related technologies, are conducted in a transparent manner and that their activities promote compliance with privacy ethics and data protection rules.

    Comment: See the earlier rule and comment number 13 about explainability. Communications should be both transparent and explainable.

28. Practitioners must ensure that the relevant data protection questions are being adequately addressed when processing personal data using AI and related technologies as part of their professional activities.

    Comment: Refer to the earlier *AI and UK Data Protection Legislation* section for more guidance.

29. Practitioners must complete a risk assessment before undertaking any data collection using AI and related technologies. Practitioners must complete a Data Protection Impact Assessments (DPIA) for high-risk AI activities where personal data is collected using AI and related technologies.

    Comment: See MRS Standards & Policy Team webinar – 'Data protection Impact Assessment (DPIA) and Data Breach Reporting'

Data Security

30. Practitioners must implement adequate technical measures to ensure that AI and related technologies include appropriate data security arrangements to minimise security risks, provide robust protection to reduce the likelihood of:

    - data breaches
    - inappropriate use of client and/or participant data
    - personal attacks and/or harm to participants.

Data Deletion

31. Practitioners must take reasonable action, including technology providers, to ensure that all participant requests for data deletion and/or data correction from AI and related technologies are undertaken appropriately.

    Comment: Data deletion can be challenging with some AI and related technologies, particularly LLMs. Practitioners should seek assurances from AI and related technology providers that any data deletion and/or data correction requests have been fulfilled, including details of the process that has been completed.

Reporting

32. Practitioners must allow clients to arrange checks on the quality of outputs produced by AI and related technologies.

33. Practitioners must record and communicate to clients, which research outputs are generated by AI and related technologies (including synthetic data) and those due to human insight.

34. Practitioners must be able to demonstrate the validity of research insights, conclusions and recommendations generated using AI and related technologies.

    Comment: This is required to avoid occurrences of activities such as AI hallucinations, a phenomenon whereby AI and related technologies generate data that is either non-existent or imperceptible to human observers in order to provide answers to questions being posed resulting in inaccurate and/or nonsensical outputs.

    The best way to mitigate the impact of AI generated hallucinations is to create systems that are less likely to hallucinate. Steps include using high quality data, using standard templates with limited responses, have a clear definition of purpose, undertake regular testing and continuous improvement and ensure there is human oversight.

35. Practitioners must provide clients with sufficient information to enable clients to assess the validity of results carried out on their behalf using AI and related technologies.

    Comment: See the MRS Code of Conduct for more details about reporting requirements.

Resources

For more guidance and advice about AI see the following links.

AI and Data Protection

- ICO overview of how to apply the principles of the UK GDPR to the use of information in AI systems: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/
- ICO and Alan Turing Institute guidance explaining decisions made with AI guidance: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/
- ICO and data protection risk toolkit: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/ai-and-data-protection-risk-toolkit/
- ICO data analytics toolkit: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/toolkit-for-organisations-considering-using-data-analytics/

EU and AI

- European AI Office: https://digital-strategy.ec.europa.eu/en/policies/ai-office
- European Commission AI innovation package for start-ups and SMEs: https://ec.europa.eu/commission/presscorner/detail/en/ip_24_383

AI and Ethics

- Centre for Data Ethics and Innovation (CDEI) AI assurance techniques: https://www.gov.uk/ai-assurance-techniques

AI and Standards

- AI Standards Hub: https://aistandardshub.org/

AI and Data

- IBM: What are AI hallucinations?: https://www.ibm.com/think/topics/ai-hallucinations
- UK Statistics Authority: Ethical considerations realting to the creastion and use of synthetic data