Advanced Certificate in Market & Social Research Practice

AI Glossary for Research & Insight Professionals

Introduction

AI is rapidly reshaping research and insight, from questionnaire design and data collection to analysis, reporting and decision-making. These tools can improve efficiency and deepen insight, while introducing new terms, concepts and ethical considerations for researchers to understand.

As AI develops, terms such as *large language models*, *agentic AI*, *synthetic data* and *retrieval-augmented generation* are becoming common in research discussions, platforms and client conversations, but are often explained in overly technical language.

This glossary offers a practical, accessible guide to key AI terms shaping research and insight. It is designed to build shared understanding for those new to AI and those already using it in their work.

It focuses on terms relevant to data collection, analysis, knowledge discovery, automation, forecasting and insight generation, as well as quality, ethics, transparency and responsible AI.

Because AI terminology continues to evolve, this glossary should be treated as a living resource that will develop alongside the field. If you have any suggestions for additions or changes, please contact samantha.driscoll@mrs.org.uk.

AI should enhance, not replace, researchers' expertise. Asking meaningful questions, applying critical thinking, ensuring rigour and communicating evidence-based insights remain essential human skills.

We hope this glossary strengthens AI literacy, supports informed professional conversations and helps researchers use AI ethically and effectively.

A

Agentic AI

AI systems designed to perform tasks autonomously by planning, making decisions, using tools and adapting their actions to achieve a goal with minimal human intervention.

Algorithm

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

Artificial Intelligence (AI)

The use of computer algorithms and systems to simulate human intelligence and perform tasks such as data analysis or decision making. See Machine Learning, which is a major subset of AI. [MRS Glossary]

AI Alignment

The practice of ensuring AI systems behave in ways that reflect human values, intentions and objectives.

AI Fluency

The ability to understand, critically evaluate and work effectively with AI technologies while recognising their strengths, limitations and ethical implications.

B

Bias

Systematic error or distortion in data, algorithms or research design that can lead to unfair, inaccurate or misleading conclusions. In AI systems, bias may arise from training data, model design or human assumptions.

Big Data

Large and complex datasets characterised by high volume, variety and velocity that can be analysed to uncover patterns, trends and relationships.

C

Computer Vision

A branch of AI that enables computers to interpret and analyse images and video. In research, it can be used for shopper observation, emotion analysis and visual content classification.

Context Window

The amount of information an AI model can consider at one time when generating responses. Larger context windows allow models to process longer documents and conversations.

D

Data Quality

A measure of the accuracy, completeness, consistency and reliability of data. High-quality data remains essential for both traditional research and AI-driven analysis.

Data Mining

Data mining is the process of discovering patterns and knowledge from large amounts of data, by using techniques such as statistical analysis, machine learning, and database systems. The goal of data mining in market research is to extract useful insights and make data-driven decisions that can inform marketing strategies, product development, and other business initiatives. [MRS Glossary]

Deep Learning

A subset of machine learning that uses multi-layered neural networks to recognise patterns and solve complex problems such as image recognition, language processing and prediction.

E

Embeddings

Numerical representations of words, phrases or documents that enable AI systems to identify meaning and relationships between pieces of information.

Explainable AI (XAI)

Approaches and technologies that help users understand how AI systems reach their conclusions or recommendations.

F

Foundation Model

A large AI model trained on extensive datasets that can be adapted to perform many tasks, such as summarisation, classification, content generation and analysis.

G

Generative AI

AI systems capable of creating new content, including text, images, audio, video and code, based on patterns learned from existing data.

Generative AI for Research

A type of AI that can 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.

Grounding

The process of connecting AI outputs to trusted sources of information, reducing the risk of inaccurate or fabricated responses.

H

Hallucination

An AI-generated response that appears plausible but is inaccurate, misleading or unsupported by evidence. Hallucinations are a recognised limitation of generative AI systems.

Human-in-the-Loop

An approach where humans review, validate or guide AI outputs to improve quality, accuracy and accountability.

I

Inference

The stage at which a trained AI model applies what it has learned to generate predictions, classifications or responses to new inputs.

L

Large Language Model (LLM)

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. [MRS: Guidance on Using AI and Related Technologies]

M

Machine Learning

Machine Learning is a subset of AI that involves training computer models to learn from data and make predictions or decisions. Machine learning is increasingly being used in market research in

helping to automate much of the classification and analysis of both structured and unstructured data. [MRS Glossary]

Model Training

The process of teaching an AI model to recognise patterns and relationships using a dataset.

Multimodal AI

AI systems capable of processing and generating multiple forms of information, such as text, images, audio and video simultaneously.

N

Natural Language Processing (NLP)

A field of AI that involves understanding and generating human language, such as voice recognition, text generation and sentiment analysis. [MRS Glossary]

Neural Network

A computing architecture inspired by the structure of the human brain that forms the foundation of many modern AI systems.

P

Predictive Analytics

Predictive analytics is the use of data, statistical algorithms, and machine learning techniques to identify the likelihood of future outcomes based on historical data. [MRS Glossary]

Prompt

The instruction, question or input provided to an AI system.

Prompt Engineering

The practice of designing and refining prompts to improve the quality, accuracy and usefulness of AI-generated outputs.

R

Reliability

The consistency of research results under similar conditions.

Research Ethics

The principles and standards governing responsible research practice, including participant welfare, consent, transparency, privacy and fairness.

Responsible AI

The development and use of AI in ways that are ethical, transparent, accountable, fair and safe.

Retrieval-Augmented Generation (RAG)

An approach that combines generative AI with information retrieval from trusted sources, allowing models to provide responses that are better grounded in evidence.

S

Sentiment Analysis

Looks at the emotion expressed in a text. It's a form of text analytics that uses natural language processing (NLP) and machine learning. It is commonly used to analyse survey responses, user/customer feedback and product reviews. Sentiment analysis is used to determine whether a given text contains negative, positive, or neutral emotions. [MRS Glossary]

Synthetic Data

Artificially generated data designed to replicate the statistical characteristics of real-world 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, while protecting privacy and reducing disclosure risks.

Synthetic Panel

An AI-generated population of virtual respondents designed to simulate the attitudes, behaviours and demographics of real consumers for exploratory analysis and testing.

Synthetic Participant

Artificially created data records designed to simulate real participants using techniques such as generative AI. They can supplement an existing human sample to fill quotas or increase sample size, or they can simulate an entire survey independently.

Note: The use of Synthetic Participants is quickly evolving. As such, this definition will continue to evolve. It is recommended that practitioners do their due diligence with regards to the usage of such data. [GDQ Glossary]

T

Training Data

The information used to train an AI model and shape its knowledge, capabilities and behaviour.

Transformer Model

A deep learning architecture that underpins modern large language models and enables efficient processing of language and context.

V

Validity

The degree to which a research method, measure or AI-assisted analysis accurately captures what it is intended to measure.

X

Explainability

The extent to which an AI system's outputs and decision-making processes can be understood and interpreted by humans.

Essential AI Terms Every Researcher Should Know

1. Artificial Intelligence (AI)
2. Generative AI
3. Large Language Model (LLM)
4. Machine Learning
5. Natural Language Processing (NLP)
6. Prompt Engineering
7. Hallucination
8. Bias
9. Agentic AI
10. Retrieval-Augmented Generation (RAG)
11. Synthetic Data
12. Human-in-the-Loop
13. Grounding
14. Explainable AI
15. Research Ethics

This condensed list provides a solid foundation for understanding how AI is changing research design, data collection, analysis and insight generation.