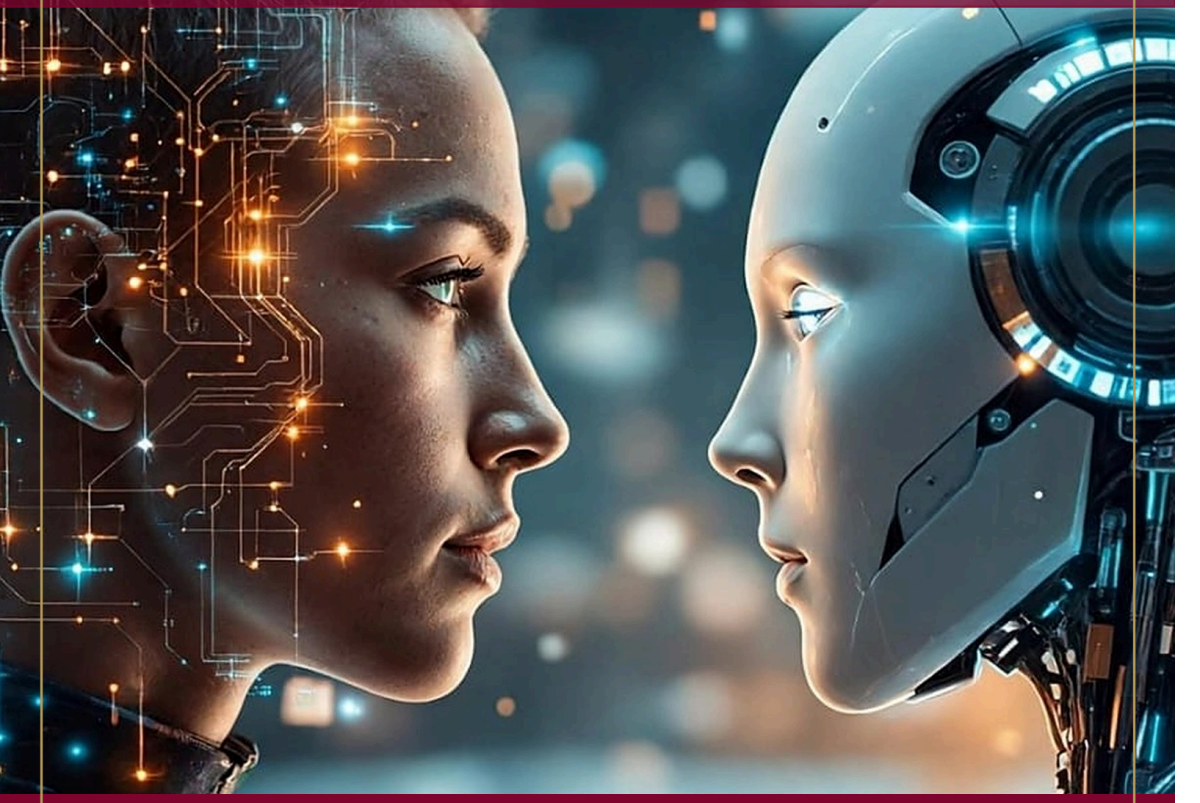
THE WYTEROCK AI COLLECTION

for doctors and clinicians



Understanding AI Terminology

Basic

Level 1: the foundations, explained in plain language

WyteRock | **CentiAgent**
your AI co-pilot

Volume 85 · WyteRock Global · wyterock.com

Understanding AI Terminology: Basic

Level 1: the foundations, explained in plain language

The WyteRock AI Collection — Volume 85

Published by WyteRock Global · https://wyterock.com

Foreword

Every new field arrives with its own vocabulary. Medicine has one, law has one, and so does accounting. Artificial intelligence now has one too, and it is spreading quickly into meetings, job descriptions, news reports and the software we use every day. Words such as *model*, *prompt*, *token*, *hallucination* and *agent* are used as if everyone already knows them. Many people nod along without being sure.

This book is for anyone who has felt that way. It is written for clinicians, managers, students, administrators, carers and curious readers. You do not need any mathematics, any programming or any technical background. If you can read a recipe, you can read this book.

It is the first of three levels. **Level 1 (Basic)**, which you are reading, explains around fifty core terms in plain language. **Level 2 (Intermediate)** and **Level 3 (Advanced)** go further, for readers who want to evaluate, buy, build or govern AI systems.

Each term is explained in the same four steps:

- **What it means**: a one-sentence definition.
- **Think of it as**: an everyday analogy.
- **In practice**: an example from healthcare or the workplace.
- **Common confusion**: the mistake people most often make with the word.

Each chapter ends with a short summary ("In one minute"), three takeaways, and a five-question self-check. The answers are at the end of the book. A glossary collects every term in one place.

Our house principle runs throughout. **You are the pilot. CentiAgent is your co-pilot.** CentiAgent is the AI co-pilot that WyteRock deploys. In this book it appears in a few illustrated screens, always showing its sources, its confidence and its limits, and always leaving the decision to a person.

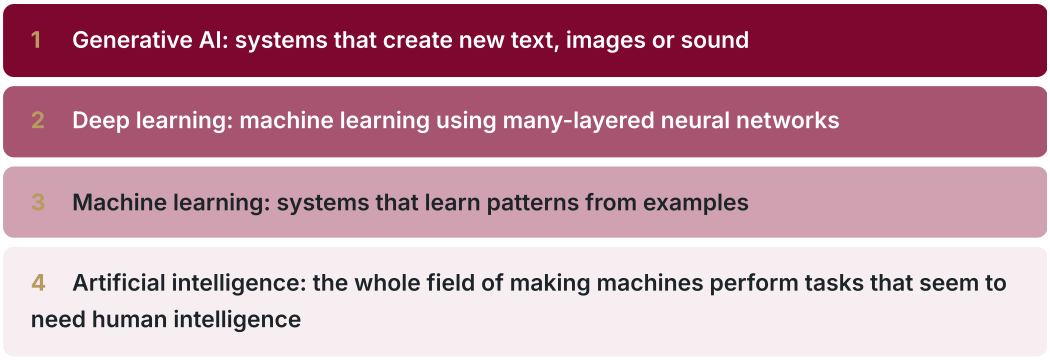
This book is educational; it is not legal, regulatory or medical advice. Every date and legal provision has been checked against a primary source, listed at the end. All people, patients and organisations in the examples are fictional.

Welcome. By the last page, the vocabulary of AI will feel like your own.

Chapter 1. The big picture: what "AI" actually means

Before the details, a map. Most confusion about AI comes from a few large words being used as if they meant the same thing. They do not. They nest inside one another, like a set of Russian dolls.

FIGURE 1. THE NESTING DOLLS OF AI



Each layer sits inside the one beneath it: all generative AI is deep learning, all deep learning is machine learning, and all machine learning is AI, but not the other way round.

Artificial intelligence (AI)

What it means. Artificial intelligence is the broad field of building computer systems that perform tasks we normally associate with human thinking, such as recognising, predicting, deciding or writing.

Think of it as the word "transport". It covers bicycles, buses and aircraft. Saying a product "uses AI" tells you about as much as saying a journey "uses transport".

In practice. A hospital scheduling tool that predicts which patients are likely to miss appointments, a phone that recognises your face, and a chatbot that drafts a letter are all described as AI, yet they work in very different ways.

Common confusion. AI is not one thing, and it is not a mind. Many people imagine a single thinking machine. In reality, AI is a large family of techniques, most of them narrow and specialised.

Augmented intelligence

What it means. Augmented intelligence is a way of describing AI that supports and extends human judgement rather than replacing it.

Think of it as a pair of reading glasses: they help you see better, but you are still the one reading.

In practice. A radiologist uses a tool that highlights areas of an X-ray worth a second look. The radiologist still reads the whole image and writes the report.

Common confusion. "Augmented" does not mean "weaker". It describes who is in charge, not how capable the tool is.

Machine learning (ML)

What it means. Machine learning is the part of AI in which a computer learns patterns from examples, instead of being given every rule by a programmer.

Think of it as learning to recognise ripe fruit by handling hundreds of peaches, rather than memorising a written rulebook about colour and firmness.

In practice. A system shown thousands of past invoices, each labelled "correct" or "error", learns to flag new invoices that look like past errors.

Common confusion. "Learning" here is not understanding. The system finds statistical patterns; it does not know why they exist.

Neural network

What it means. A neural network is a type of machine-learning system made of many simple connected units, arranged in layers, that pass signals to one another and adjust those connections as they learn.

Think of it as a large team passing notes along rows of desks. Each person makes a small judgement and passes it on; the final row gives the answer. Over time, the team learns whose notes to trust more.

In practice. Neural networks power speech recognition in dictation software and image recognition in skin-lesion triage apps.

Common confusion. Neural networks were loosely inspired by the brain, but they are not artificial brains. The resemblance is mostly in the name.

Deep learning

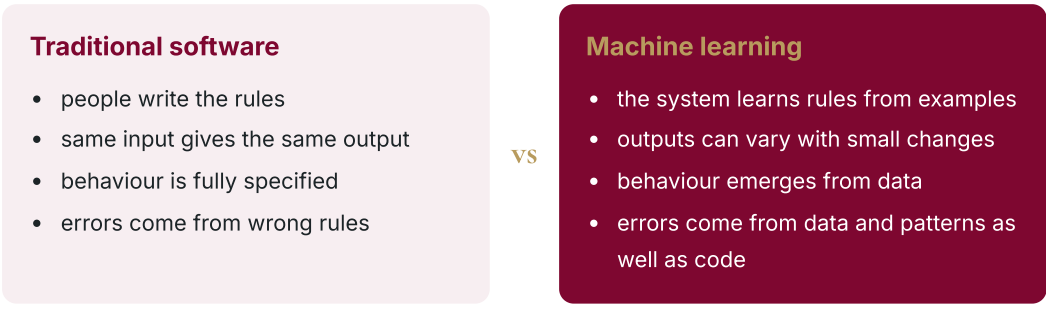
What it means. Deep learning is machine learning that uses neural networks with many layers, which lets them learn very complex patterns from large amounts of data.

Think of it as an assembly line with many stations. Early stations notice simple things, such as edges in a picture; later stations combine them into shapes, then objects.

In practice. Deep learning lies behind modern image analysis of retinal photographs and behind the language tools that draft text.

Common confusion. "Deep" refers to the number of layers, not to depth of thought or insight.

FIGURE 2. TRADITIONAL SOFTWARE VERSUS MACHINE LEARNING



Traditional software follows rules people wrote; machine learning infers its own rules from examples, which is why it needs different checks.

In one minute

- AI is a broad field, not a single machine or mind.
- Machine learning learns from examples; deep learning uses many-layered neural networks; generative AI creates new content. Each sits inside the one before.
- All AI in use today is narrow, however fluent it seems.

Three practical takeaways

1. When someone says a product "uses AI", ask which kind, and for which task.
2. Treat "learning" as pattern-finding, not understanding.
3. Use "augmented intelligence" to keep the human role explicit.

Self-check quiz, Chapter 1

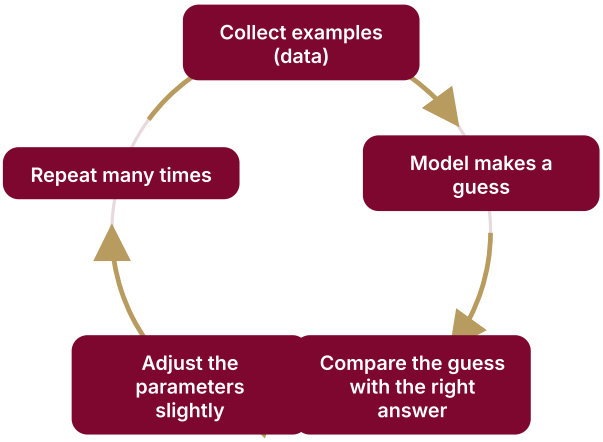
1. Which is the broadest term: deep learning, AI or machine learning?

2. True or false: neural networks are artificial brains.
3. What does "deep" mean in "deep learning"?
4. Does "augmented intelligence" mean a weaker kind of AI?
5. What is the main difference between traditional software and machine learning?

Chapter 2. How AI learns: data, training and models

If AI learns from examples, the examples matter enormously. This chapter explains the words for the ingredients and the cooking.

FIGURE 3. HOW A MODEL LEARNS



Training is a loop of guessing, checking and adjusting, repeated until the model's guesses are good enough on examples it has seen.

Dataset

What it means. A dataset is an organised collection of data gathered for a particular purpose, such as training or testing a system.

Think of it as a recipe book's tested set of ingredients, weighed and labelled, rather than whatever happens to be in the cupboard.

In practice. A dataset of 10,000 anonymised chest X-rays (fictional), each with the final diagnosis attached, could be used to build a triage tool.

Common confusion. A dataset is a snapshot. It reflects the people, places and time it came from, and may not match your own patients or customers.

Training

What it means. Training is the process of showing a machine-learning system many examples so that it gradually adjusts itself to make better predictions.

Think of it as practising scales on the piano: repetition with feedback until the fingers get it right.

In practice. A vendor trains a sepsis-alert model on years of historical hospital records before it is ever offered to a hospital.

Common confusion. Most tools you use are not learning from you in real time. Training usually happens beforehand, by the developer. Whether your inputs are used for future training depends on the product's terms and settings, which is worth checking.

Model

What it means. A model is the result of training: the learned system that takes an input and produces an output, such as a prediction, a classification or a piece of text.

Think of it as the experienced cook who emerges from years of practice. The practice was the training; the cook is the model.

In practice. "The model flagged this patient as high risk" means the trained system produced that output from the patient's data.

Common confusion. A model is not the same as the app. An app is the whole product you see; the model is the engine inside it. One app can use several models.

Parameters

What it means. Parameters are the internal settings a model adjusts during training; together they store what the model has learned.

Think of it as the thousands of small dials on a sound-mixing desk, each turned slightly until the music sounds right.

In practice. When news reports say a language model has "billions of parameters", they mean billions of these learned dials.

Common confusion. More parameters do not guarantee a better or safer tool for your task. Size is one factor among many, including the quality of training data and how the tool is tested.

Training data and test data

What it means. Training data is what a model learns from; test data is a separate set, kept aside, used to check how well it performs on examples it has never seen.

Think of it as revising from past papers, then sitting a real exam with new questions. Scoring well on the past papers alone proves little.

In practice. A developer reports results on a test set drawn from different hospitals to show the tool works beyond the place it was built.

Common confusion. Strong results on test data from one setting do not guarantee the same results in yours. Different populations, equipment and workflows can change performance.

Overfitting

What it means. Overfitting happens when a model learns its training examples too closely, including their quirks, so it performs worse on new cases.

Think of it as a student who memorises past exam answers word for word and is lost when a question is phrased differently.

In practice. A model that learned to spot pneumonia partly from a hospital's scanner markings, rather than from the lungs, would fail at another hospital.

Common confusion. A model that is "perfect" on its training data is a warning sign, not a triumph.

Inference

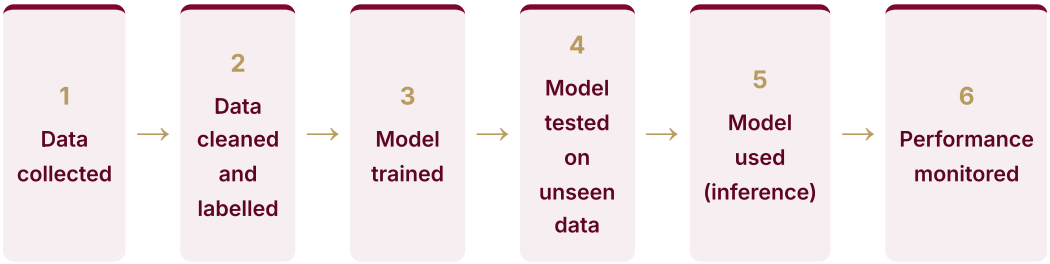
What it means. Inference is the moment a trained model is used: it receives a new input and produces an output.

Think of it as the cook preparing tonight's dinner, as opposed to the years spent learning to cook.

In practice. Each time you ask a chatbot a question, the model performs inference.

Common confusion. In AI, "inference" simply means "using the model", not drawing a careful logical conclusion as in everyday speech.

FIGURE 4. FROM DATA TO A WORKING TOOL



A trustworthy tool passes through every stage; skipping testing or monitoring is where most problems begin.

In one minute

- Data is the ingredient; training is the practice; the model is the result.
- Parameters are the learned dials inside a model; more is not automatically better.
- Always ask whether a tool was tested on data like yours.

Three practical takeaways

1. Ask any vendor: "What data was this trained on, and how was it tested?"
2. Remember that labels and data carry human mistakes.
3. Separate the model (the engine) from the app (the car).

Self-check quiz, Chapter 2

1. What is the difference between training data and test data?
2. In AI, what does "inference" mean?
3. Why can a model that scores perfectly on training data be a concern?
4. What are parameters, in one sentence?
5. True or false: most AI tools you use are learning from you in real time.

Chapter 3. Generative AI and large language models

Generative AI is the part of AI most people meet first, usually as a chatbot. This chapter explains how it handles words.

Generative AI

What it means. Generative AI is AI that creates new content, such as text, images, audio or code, in response to an instruction.

Think of it as an extremely well-read improviser: it has absorbed a vast amount of material and produces something new in a similar style.

In practice. Drafting a patient information leaflet, summarising meeting notes, or generating an illustration for a training slide.

Common confusion. "New" does not mean "verified". Generative AI produces plausible content; checking whether it is correct remains your job.

Large language model (LLM)

What it means. A large language model is a deep-learning model trained on very large amounts of text to predict which words are likely to come next, which lets it write, summarise, translate and answer questions.

Think of it as the autocomplete on your phone, scaled up enormously, so that it can continue not just a word but a whole letter or report.

In practice. Chat assistants used in offices and clinics run on LLMs. A clinician might ask one to turn bullet notes into a referral letter draft.

Common confusion. An LLM is not a database of facts and does not look things up by default. It generates text from patterns, which is why it can be fluent and wrong at the same time.

Token

What it means. A token is the small chunk of text, often a word or part of a word, that a language model reads and writes one at a time.

Think of it as the pieces of a jigsaw. The model does not see whole sentences; it sees pieces and works out which piece comes next.

In practice. A long word such as "pharmacovigilance" may be split into several tokens, while "the" is usually one. AI services often set limits and prices by the number of tokens.

Common confusion. A token is not exactly a word. Counts of tokens and counts of words differ, and the difference varies by language.

FIGURE 5. HOW TEXT BECOMES TOKENS AND BACK



A language model writes its reply one small piece at a time, each chosen as a likely continuation of what came before.

Context window

What it means. The context window is the maximum amount of text, measured in tokens, that a model can take into account at once, including your instructions, any documents and its own reply.

Think of it as the size of a desk. Whatever fits on the desk can be consulted; whatever falls off is no longer in view.

In practice. If you paste a very long policy document and the tool ignores the last section, the document may have exceeded the context window, or the tool may simply have given it less attention.

Common confusion. A large context window is not memory. When a conversation ends, most tools do not retain it unless a separate memory feature is switched on.

Prompt

What it means. A prompt is the instruction or question you give to a generative AI tool.

Think of it as a briefing to a capable new colleague: the clearer the brief, the better the work.

In practice. "Summarise this discharge letter for the patient in plain English, under 150 words, and list the medicines separately" is a clear prompt.

Common confusion. A prompt is not a magic formula. There is no secret phrase; clarity about the task, the audience and the format matters most.

Prompt engineering

What it means. Prompt engineering is the practice of writing and refining prompts to get more useful and reliable outputs.

Think of it as learning to give good directions: landmarks, distances and the destination, not just "go that way".

In practice. A practice manager adds an example of a good reply to her prompt, and the tool's draft emails become more consistent.

Common confusion. It sounds technical but is mostly good communication. Anyone who writes clear instructions already has the core skill.

Fine-tuning

What it means. Fine-tuning is additional training of an existing model on a smaller, specialised dataset so that it performs better at a particular task.

Think of it as a general practitioner completing a short course in dermatology.

In practice. A vendor fine-tunes a general model on anonymised clinic letters so its drafts follow local style.

Common confusion. Fine-tuning is different from prompting. Prompting changes the instruction; fine-tuning changes the model itself.

Retrieval-augmented generation (RAG)

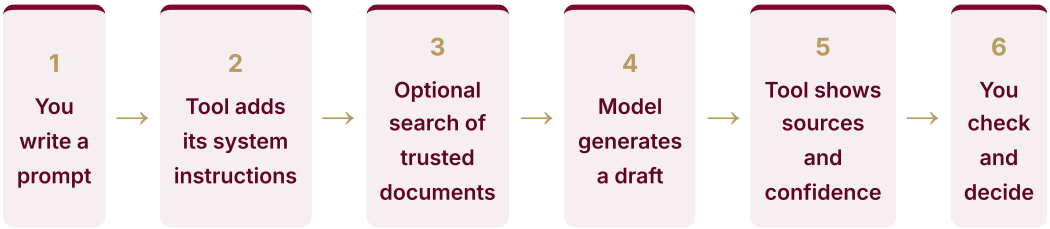
What it means. Retrieval-augmented generation is a method in which the tool first searches a trusted set of documents and then writes its answer using what it found.

Think of it as an open-book exam, rather than answering from memory.

In practice. A staff assistant answers "What is our policy on annual leave carry-over?" by finding the current HR policy and quoting the relevant paragraph.

Common confusion. RAG reduces made-up answers but does not eliminate them. The tool can still misread or misquote a source, so the source should be shown and checked.

FIGURE 6. FROM PROMPT TO ANSWER



The final step is always human: the answer is a draft until a person has checked it against its sources.

SCREEN 1. CENTIAGENT: EXPLAIN A TERM

CentiAgent: explain a term — Example screen · fictional patient

User: "Sam" (fictional), practice manager, AI-literacy module 3

YOUR QUESTION

- "In the vendor's brochure it says the tool has a 'large context window'. What does that mean for us?"

CENTIAGENT EXPLANATION

Confidence: High (definition taken from the course glossary)

The context window is how much text the tool can consider at once, measured in tokens. A large window means you can give it longer documents, such as a full policy, in one go. It does not mean the tool remembers past conversations. Ask the vendor how much text fits and whether any memory feature stores your data.

Sources

- WyteRock Level 1 glossary: "Context window"
- "Token"

⚠ The brochure's claim about speed is not covered by this explanation. Check it with the vendor.

Accept | Edit | Reject | Ask why

A good co-pilot explains a term in plain words, names its source, and points out what it has not answered.

In one minute

- Generative AI creates content; a large language model predicts text one token at a time.
- The context window is how much the model can consider at once; it is not memory.
- Clear prompts matter more than clever ones; RAG helps by grounding answers in trusted documents.

Three practical takeaways

1. Write prompts that name the task, the audience, the length and the format.
2. Prefer tools that show their sources.
3. Never assume a generative tool has checked its own facts.

Self-check quiz, Chapter 3

1. Is a token always exactly one word?
2. What is the difference between prompting and fine-tuning?
3. True or false: a large context window means the tool remembers your past conversations.
4. What does retrieval-augmented generation do first, before writing?
5. True or false: "new" content from generative AI has been verified.

Chapter 4. What comes out: outputs, hallucinations and bias

An AI tool's answer can be helpful, partly wrong or confidently invented. This chapter gives you the words to tell them apart.

Output

What it means. An output is whatever an AI system produces in response to an input: a prediction, a score, a label, a summary or a draft.

Think of it as a dish served by the kitchen. It may look excellent, but tasting it is still wise.

In practice. A risk score of "high" for a patient, a draft reply to a complaint, or a list of suggested diagnoses.

Common confusion. An output is a suggestion, not a decision. The decision belongs to the person who reviews it.

Hallucination

What it means. A hallucination is an output from a generative AI tool that sounds plausible but is false or unsupported, such as an invented fact, quotation or reference.

Think of it as a confident guest at dinner who fills a gap in a story with details that never happened, without realising it.

In practice. A tool asked for evidence cites a journal article with a real-sounding title and authors that does not exist. Or it adds a medication to a summary that was never prescribed.

Common confusion. A hallucination is not a lie: the tool has no intention. Nor is it rare or a sign of a broken product. It is a known limitation of how language models generate text, so checking is always needed.

FIGURE 7. HALLUCINATION VERSUS A GROUNDED FACT

Hallucination

- no source shown or source cannot be found
- specific details that cannot be traced
- same confidence as true statements
- changes when you ask again

VS

Grounded fact

- a named, checkable source
- wording matches the source
- uncertainty stated where it exists
- consistent when checked again

Fluency is not evidence; a statement you can trace to a source is far more trustworthy than one you cannot.

SCREEN 2. CENTIAGENT: CLINIC LETTER CHECK

CentiAgent: clinic letter check — Example screen · fictional patient

Patient: "Mr B." (fictional), 64, outpatient cardiology follow-up

SOURCE RECORD

- Clinic note: blood pressure controlled
- continue ramipril 5 mg
- review in 6 months. Medication list: ramipril 5 mg, atorvastatin 20 mg.

CENTIAGENT DRAFT LETTER

Confidence: Medium (two statements matched to the record; one not found)

Dear Dr Patel (fictional), I reviewed Mr B. today. His blood pressure is well controlled on ramipril 5 mg and atorvastatin 20 mg, which he will continue. I have arranged a review in six months.

Sources

- Clinic note (today)
- Medication list (today)

⚠ An earlier draft mentioned "bisoprolol 2.5 mg", which does not appear in the source record. It has been removed. Please confirm the medication list.

Accept | Edit | Reject | Ask why

A hallucination flag compares the draft with the record and shows the clinician exactly what could not be traced.

Bias

What it means. Bias in AI is a systematic tendency for a system to produce results that are unfair or less accurate for certain groups, often because of the data it learned from.

Think of it as a set of bathroom scales that always reads two kilos light for some people and not for others.

In practice. A skin-condition tool trained mostly on images of lighter skin may perform less well on darker skin. A recruitment tool trained on past hiring may repeat past patterns.

Common confusion. Bias is not only about intention or prejudice. It is often an unnoticed gap in the data, which is why it must be tested for, not assumed absent.

Drift

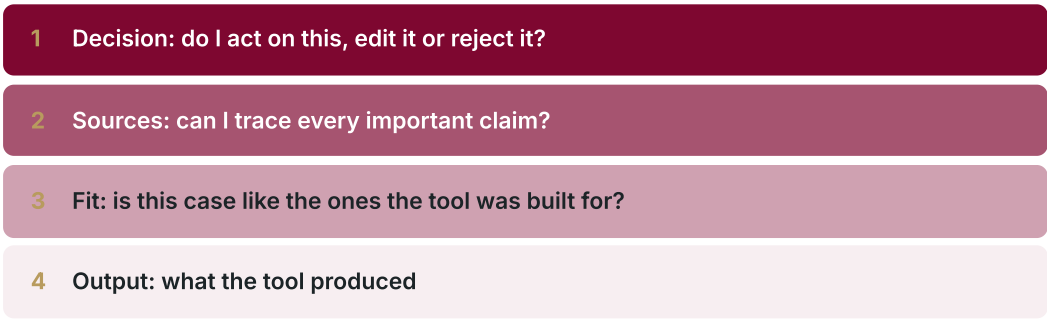
What it means. Drift is the gradual decline in a model's performance as the real world changes away from the data it was trained on.

Think of it as a printed map slowly going out of date as new roads are built.

In practice. A model trained before a new laboratory analyser was installed may behave differently once results are reported in a slightly different way.

Common confusion. A tool that worked well at launch is not guaranteed to work well a year later. Monitoring is part of using AI, not an optional extra.

FIGURE 8. THREE CHECKS BEFORE YOU TRUST AN OUTPUT



Read upwards from the output; trust grows only as each check is passed.

In one minute

- An output is a suggestion; a hallucination is a plausible but false output.

- Bias is a systematic gap, often from the data; drift is decline over time.
- Confidence labels help, but do not replace checking.

Three practical takeaways

1. Trace every important fact in an AI draft to its source.
2. Ask how a tool performs for different groups of people.
3. Expect performance to change over time, and ask who monitors it.

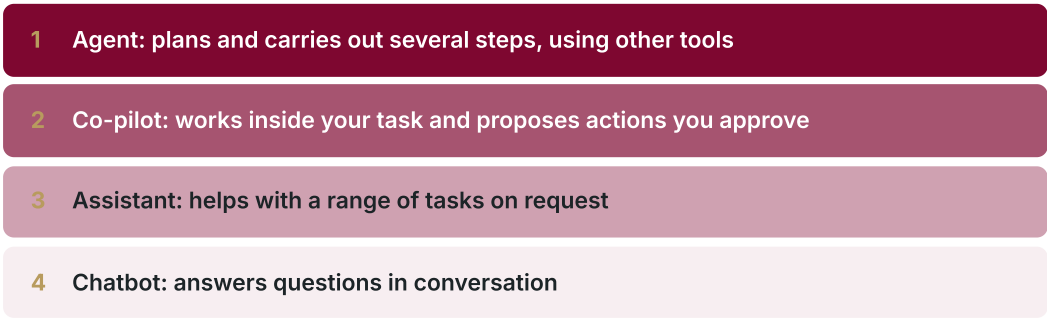
Self-check quiz, Chapter 4

1. Is a hallucination a deliberate lie?
2. Give one common cause of bias in an AI system.
3. What is drift?
4. True or false: a tool that worked well at launch will keep working well.
5. What is the single most useful habit for catching hallucinations?

Chapter 5. Chatbots, assistants, agents and co-pilots

These four words are used loosely, sometimes interchangeably. They describe different degrees of help and of independence.

FIGURE 9. FROM ANSWERING TO ACTING



Moving up the ladder, tools do more on their own, so the need for clear limits and human approval grows.

Chatbot

What it means. A chatbot is a program that holds a written or spoken conversation with a person.

Think of it as an information desk that answers questions.

In practice. A clinic website's chatbot answers "What are your opening hours?" or "How do I book a blood test?"

Common confusion. Not every chatbot uses generative AI. Many older ones follow fixed scripts and menus.

AI assistant

What it means. An AI assistant is a tool, usually based on a large language model, that helps with a range of tasks on request, such as writing, summarising and answering questions.

Think of it as a capable general helper who waits to be asked.

In practice. A manager asks an assistant to summarise a long email thread and draft a reply.

Common confusion. An assistant is not an authority. It is as good as its sources and as limited as its training.

Co-pilot

What it means. A co-pilot is an AI tool built into your work that suggests, drafts and checks, while you keep the final say.

Think of it as the co-pilot of an aircraft: skilled, active and helpful, but the captain remains responsible.

In practice. CentiAgent drafts a discharge summary from the record, flags missing allergies, and waits for the clinician to accept, edit or reject it.

Common confusion. "Co-pilot" is also a product name used by some companies. In this book it describes a way of working: the human is the pilot.

Agent (agentic AI)

What it means. An AI agent is a system that can plan and carry out a sequence of steps towards a goal, often using other software tools, with less step-by-step instruction from a person.

Think of it as a trusted errand-runner who is told "organise the team's training day" and then books the room, checks diaries and sends invitations.

In practice. An agent receives a referral, checks the clinic's availability, proposes an appointment and drafts the confirmation letter, pausing for staff approval before sending.

Common confusion. "Agent" does not mean fully independent. Well-designed agents have limits, logs and approval points. The more an agent can do, the more important those controls become.

Human in the loop

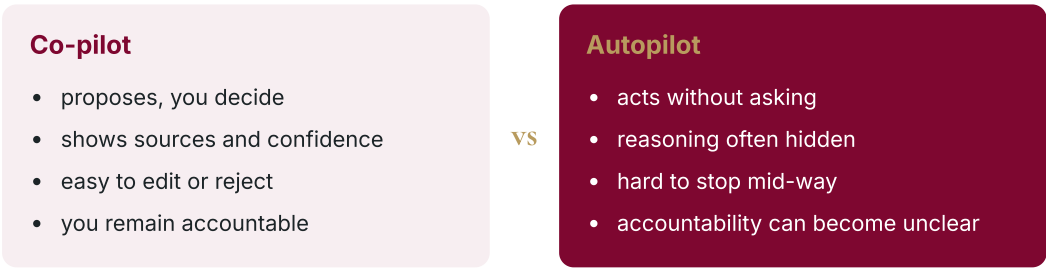
What it means. Human in the loop means a person reviews, approves or corrects an AI system's output before it has an effect.

Think of it as a second signature on an important cheque.

In practice. An AI-drafted letter to a patient is never sent until a clinician has read and approved it.

Common confusion. A human who clicks "approve" without reading is not truly in the loop. Meaningful review needs time, information and authority to say no.

FIGURE 10. CO-PILOT VERSUS AUTOPILOT



In health and in most workplaces, the co-pilot model keeps responsibility where it belongs: with a person.

In one minute

- Chatbots converse; assistants help on request; co-pilots work inside your task; agents carry out multi-step actions.
- More independence needs more limits and clearer approval points.
- Human in the loop only works when the human genuinely reviews.

Three practical takeaways

1. For any tool, ask: "What can it do without my approval?"
2. Insist on visible sources and an easy way to reject.
3. Allow time for real review; a rushed click is not oversight.

Self-check quiz, Chapter 5

1. Does every chatbot use generative AI?
2. What is the key difference between an assistant and an agent?
3. In the co-pilot model, who is accountable for the decision?
4. Can automation work without AI?
5. When is a "human in the loop" not really in the loop?

Chapter 6. Measuring performance without the maths

How do we know whether an AI tool is any good? Several words are used, and two of them come straight from medicine. No formulas are needed; the ideas are what matter.

True and false positives and negatives

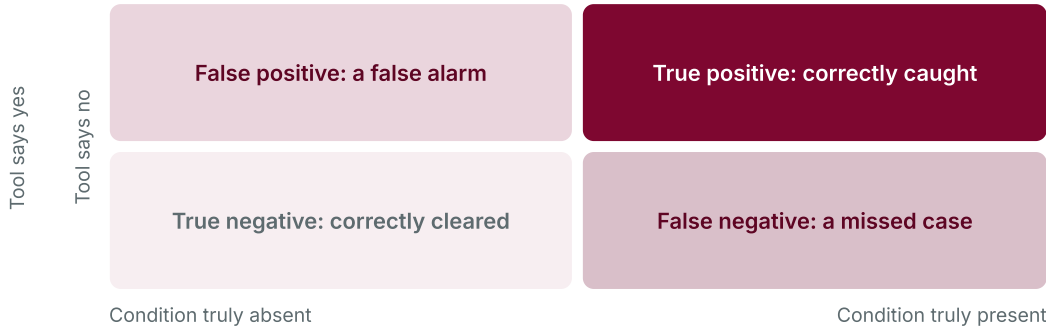
What it means. When a tool says "yes" or "no" about something, there are four possible results: it correctly says yes (true positive), wrongly says yes (false positive), correctly says no (true negative) or wrongly says no (false negative).

Think of it as a smoke alarm. It can ring for a real fire, ring for burnt toast, stay silent when all is well, or stay silent during a real fire.

In practice. A screening tool flags a mammogram that does show cancer (true positive), or flags one that is normal (false positive, causing worry and extra tests).

Common confusion. The two kinds of error do not cost the same. A missed fire is far worse than a false alarm, and the right balance depends on the situation.

FIGURE 11. THE FOUR POSSIBLE RESULTS OF A YES-OR-NO TOOL



Every yes-or-no tool lands each case in one of these four boxes; sensitivity and specificity describe how it handles the columns.

Sensitivity

What it means. Sensitivity is how good a tool is at catching the cases that really are there; out of all the people who truly have the condition, what share does it correctly flag.

Think of it as a fishing net with a fine mesh: it catches nearly every fish.

In practice. A highly sensitive sepsis alert rarely misses a patient with sepsis, but may also alert on many who do not have it.

Common confusion. High sensitivity alone says nothing about false alarms. A tool that flags everyone has perfect sensitivity and is useless.

Specificity

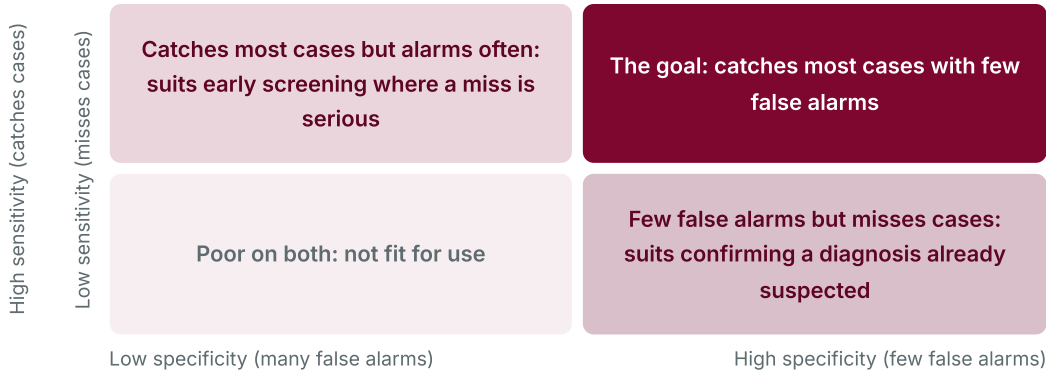
What it means. Specificity is how good a tool is at correctly clearing the cases that are not there; out of all the people who truly do not have the condition, what share does it correctly leave unflagged.

Think of it as a careful bouncer who lets in everyone on the guest list and does not turn away invited guests by mistake.

In practice. A highly specific test rarely labels a healthy person as ill, which spares unnecessary worry and procedures.

Common confusion. Sensitivity and specificity usually pull against each other. Making a tool catch more real cases often raises false alarms, and the reverse.

FIGURE 12. CHOOSING THE BALANCE BETWEEN SENSITIVITY AND SPECIFICITY



The right balance depends on the purpose; screening usually favours sensitivity, confirmation usually favours specificity.

Accuracy (the measured figure)

What it means. As a measurement, accuracy is the share of all cases that a tool gets right, adding together correct yeses and correct noes.

Think of it as a single overall exam mark.

In practice. If a condition affects very few people, a tool that always says "no" will be right almost every time, giving a high accuracy figure while missing every real case.

Common confusion. High accuracy can hide poor sensitivity when the condition is rare. Always ask for sensitivity and specificity as well.

False alarm and alert fatigue

What it means. Alert fatigue is what happens when people receive so many alerts, many of them false alarms, that they start to ignore or dismiss them.

Think of it as the car alarm in the street that everyone has stopped noticing.

In practice. Staff on a busy ward begin dismissing a deterioration alert without reading it, because most of its alerts have not been meaningful.

Common confusion. A tool that alerts more is not necessarily safer. Too many false alarms can make real warnings easier to miss.

In one minute

- Every yes-or-no tool produces true and false positives and negatives.
- Sensitivity is catching real cases; specificity is correctly clearing non-cases; they usually trade off.
- Accuracy alone can mislead, especially for rare conditions.

Three practical takeaways

1. Ask for sensitivity and specificity, not just accuracy.
2. Decide which error is worse for your purpose before choosing a tool.
3. Watch for alert fatigue after any tool goes live.

Self-check quiz, Chapter 6

1. What is a false negative, in plain words?
2. A tool flags every patient as positive. What is its sensitivity, and why is it useless?
3. Which usually matters more for early screening: sensitivity or specificity?
4. Why can high accuracy mislead for a rare condition?
5. What is alert fatigue?

Chapter 7. Privacy, personal data and the GDPR basics

AI runs on data, and much of that data is about people. In Europe the rules are set mainly by the General Data Protection Regulation (GDPR), which has applied since 25 May 2018. This chapter explains the core words. It is not legal advice; your data-protection officer is the right person for specific questions.

Personal data

What it means. Under the GDPR (Article 4), personal data is any information relating to an identified or identifiable living person.

Think of it as anything that could lead back to a particular person, on its own or combined with other information.

In practice. A name, a date of birth, an NHS or hospital number, an email address, a staff ID, and even a combination such as "the only 90-year-old on Ward 3" can all be personal data.

Common confusion. Removing the name is not enough. If someone could reasonably work out who the person is, the information is still personal data.

Special category data (including health data)

What it means. Special category data is personal data that the GDPR (Article 9) protects more strictly, including data concerning health, genetic data, and biometric data used to identify a person.

Think of it as the locked drawer within an already locked filing cabinet.

In practice. Diagnoses, test results, prescriptions and clinic letters are health data. Using them, including in an AI tool, requires a specific legal basis and appropriate safeguards.

Common confusion. Health data is not forbidden to use. It can be processed lawfully, for example for providing care, but only under the specific conditions the law sets.

Anonymisation and pseudonymisation

What it means. Anonymised data can no longer be linked to a person by any reasonably likely means; pseudonymised data has identifiers replaced by codes but can still be re-linked by someone holding the key.

Think of it as shredding a name tag (anonymisation) versus putting a numbered sticker over it with the list kept in a drawer (pseudonymisation).

In practice. A research dataset where patients appear as "P-0417" with the lookup table held by the hospital is pseudonymised, and it is still personal data under the GDPR.

Common confusion. Pseudonymised data is often wrongly called "anonymous". True anonymisation is harder than it looks, especially with rich health records.

Consent and lawful basis

What it means. A lawful basis is the legal reason that allows personal data to be used; consent is one of six lawful bases under the GDPR (Article 6), not the only one.

Think of it as a set of different keys, any one of which can open the door if it fits the situation.

In practice. Much healthcare data is processed on bases other than consent, such as tasks carried out in the public interest, alongside the extra conditions for health data in Article 9.

Common confusion. "We need the patient's consent to use AI" is not always correct, and nor is "we never need consent". The right basis depends on the purpose and should be decided with your data-protection officer.

Data minimisation

What it means. Data minimisation is the GDPR principle (Article 5) that you should use only the personal data that is adequate, relevant and necessary for the purpose.

Think of it as packing only what you need for the trip.

In practice. To ask an AI tool to improve the wording of a template letter, you do not need to paste in a real patient's name and history.

Common confusion. "The tool can handle it" is not a reason to share more. The question is whether the purpose needs it.

Automated decision-making

What it means. The GDPR (Article 22) gives people the right, with certain exceptions, not to be subject to a decision based solely on automated processing that has legal or similarly significant effects on them.

Think of it as the right to have a person, not only a machine, decide something important about you.

In practice. If an AI system alone decided who receives a service, with no meaningful human review, Article 22 would need careful attention.

Common confusion. This does not ban AI support. It concerns decisions made *solely* by automated means. A person genuinely reviewing the AI's suggestion changes the picture.

Data protection impact assessment (DPIA)

What it means. A DPIA is a structured assessment, required by the GDPR (Article 35) when processing is likely to result in a high risk to people's rights, that identifies risks and how they will be reduced.

Think of it as a risk assessment before a building project, done before the first wall goes up.

In practice. Before a clinic introduces an ambient note-taking tool that records consultations, the organisation carries out a DPIA.

Common confusion. A DPIA is not paperwork after the fact. It is meant to shape the decision, including whether to go ahead.

FIGURE 13. A PRIVACY CHECK BEFORE USING AN AI TOOL



Five quick questions, asked before pasting anything into an AI tool, prevent most everyday privacy mistakes.

SCREEN 3. CENTIAGENT: PRIVACY CHECK

CentiAgent: privacy check — Example screen · fictional patient

User: "Dr C." (fictional), GP registrar, about to paste text into a drafting tool

TEXT TO BE PASTED

- "Mrs D. (fictional), DOB 14/03/1951, NHS no. 000 000 0000 (fictional), lives alone at 12 Elm Road (fictional), recent fall, on warfarin."

CENTIAGENT PRIVACY CHECK

Confidence: High (four identifiers detected by pattern; health terms detected)

This text contains a name, date of birth, health identifier and address, plus health data. For the task you described (improving the wording of a falls-advice letter), none of these identifiers is needed. I suggest replacing them with placeholders such as [Patient] and [Address] before continuing.

Sources

- Your organisation's AI use policy, section 3
- GDPR Articles 4, 5 and 9 (summary)

⚠ Pattern detection can miss unusual identifiers, such as a rare job title or a unique family detail. Read the text once more before sending.

Accept | Edit | Reject | Ask why

A privacy check suggests removing identifiers, explains why, admits what it might miss, and leaves the choice with the user.

In one minute

- Personal data is anything that can identify a living person; health data is a special category with stronger protection.
- Pseudonymised data is still personal data; true anonymisation is hard.
- Use the minimum data needed, use approved tools, and involve your data-protection officer early.

Three practical takeaways

1. Before pasting text into any AI tool, remove identifiers the task does not need.
2. Use only tools your organisation has approved for the data involved.
3. Ask whether a DPIA has been done before a new AI tool goes live.

Self-check quiz, Chapter 7

1. Is pseudonymised data still personal data under the GDPR?
2. Name two examples of special category data.
3. True or false: consent is the only lawful basis for using personal data.
4. What does data minimisation mean in one sentence?
5. What does Article 22 of the GDPR concern?

Chapter 8. The human in control, and the EU AI Act in one page

The final chapter brings everything together: who is responsible, and what the main European AI law says.

Accountability

What it means. Accountability means a named person or organisation is responsible for how an AI system is used and for the decisions made with it.

Think of it as the name on the prescription: whoever signs it answers for it.

In practice. A clinician who accepts an AI-drafted summary and signs it is accountable for its content, as with any document they sign.

Common confusion. "The AI decided" is never an acceptable explanation. Tools do not hold responsibility; people and organisations do.

Human oversight

What it means. Human oversight means people can understand, monitor, intervene in and, if necessary, stop an AI system.

Think of it as the emergency brake on a train, together with a driver who knows when to use it.

In practice. A ward manager can switch off a misbehaving alert, report the problem and return to the previous process without disruption.

Common confusion. Oversight is not the same as being present. It requires the information, training and authority to act.

AI literacy

What it means. AI literacy is the skills, knowledge and understanding that allow people to use AI systems sensibly, aware of both their opportunities and their risks.

Think of it as knowing how to read a map, not just owning one.

In practice. A team that understands hallucination, bias and data protection uses an AI assistant more safely and gets more value from it.

Common confusion. AI literacy is not coding. It is the vocabulary and judgement this book aims to give you.

Shadow AI

What it means. Shadow AI is the use of AI tools at work without the organisation's knowledge or approval.

Think of it as using a personal memory stick for work files: convenient, but outside the organisation's protections.

In practice. A staff member pastes meeting notes containing patient details into a free public chatbot on a personal account.

Common confusion. Shadow AI usually comes from good intentions: people want to work faster. The answer is approved, safe tools and training, not only prohibition.

The EU AI Act in one page

The **Artificial Intelligence Act**, Regulation (EU) 2024/1689, is the European Union's law on AI. It entered into force on 1 August 2024 and applies in stages. It was amended by **Regulation (EU) 2026/1744** (often called the Digital Omnibus on AI), published on 24 July 2026 and in force from 27 July 2026.

The idea. The Act takes a risk-based approach: the greater the risk an AI system poses to health, safety or fundamental rights, the stricter the rules.

FIGURE 14. THE RISK-BASED APPROACH OF THE EU AI ACT

1 **Prohibited practices: banned uses, such as certain manipulative or exploitative systems (Article 5)**

2 **High-risk systems: strict requirements, including many AI medical devices and listed uses (Article 6, Annexes I and III)**

3 **Transparency duties: people must be told in certain cases, for example when talking to an AI or seeing AI-generated content (Article 50)**

4 **Minimal risk: most everyday AI, with no specific new obligations beyond AI literacy**

The higher the layer, the stricter the rules; most everyday workplace tools sit in the lower layers.

Key dates in plain words.

- **2 February 2025:** the general provisions, including the AI-literacy duty (Article 4), and the bans on prohibited practices (Article 5) began to apply.
- **2 August 2026:** most remaining provisions began to apply, including the transparency duties in Article 50.
- **2 December 2027 and 2 August 2028:** after the 2026 amendments, the high-risk obligations apply from 2 December 2027 for systems listed in Annex III and from 2 August 2028 for systems covered by the product laws in Annex I, which include medical devices.

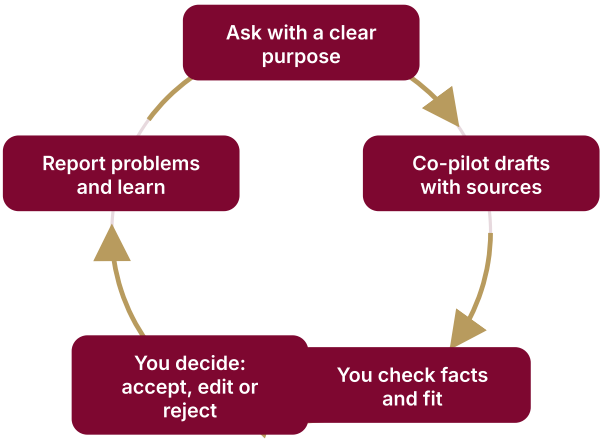
The AI-literacy duty (Article 4). Providers and deployers of AI systems must take measures to support the AI literacy of their staff and of other people using AI systems on their behalf, taking into account their knowledge, experience, education and training and the context of use. Following the 2026 amendments, the consolidated text makes clear that this does not require any specific level of AI literacy for each individual. It remains a duty to act: organisations are expected to do something sensible and proportionate.

Who is who. A **provider** develops an AI system or places it on the market. A **deployer** uses an AI system under its own authority in a professional setting. A hospital using a vendor's tool is usually a deployer; the vendor is usually the provider.

What it means for you. If your organisation uses AI, it should be supporting your AI literacy, telling people when they are dealing with AI where required, and keeping humans in genuine control of significant decisions. Reading this book is one part of that.

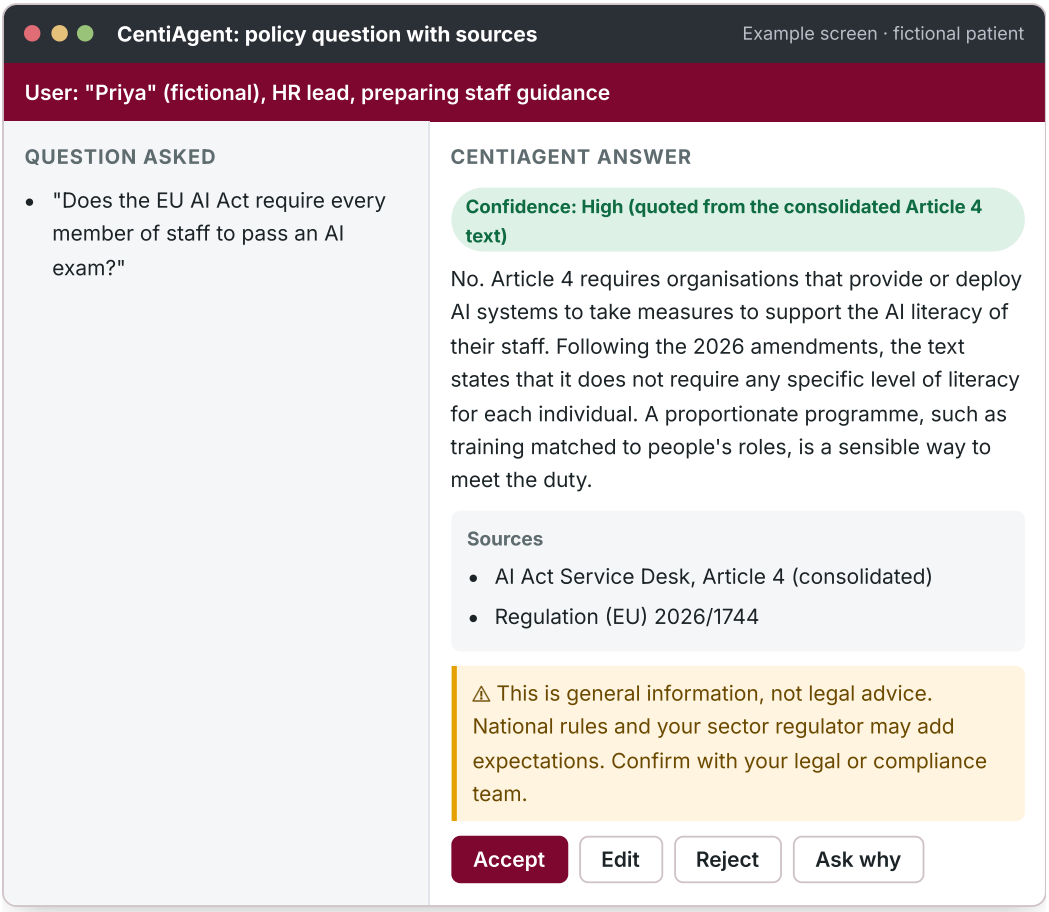
Common confusion. The AI Act does not replace the GDPR. Both apply. The AI Act governs AI systems; the GDPR governs personal data. A health AI tool will often be subject to both, and to medical-device law as well.

FIGURE 15. THE PILOT'S LOOP WITH A CO-PILOT



Each turn of the loop keeps the human as pilot, and each reported problem makes the next turn safer.

SCREEN 4. CENTIAGENT: POLICY QUESTION WITH SOURCES



A grounded answer quotes the rule, names its source, and is honest about where its knowledge stops.

In one minute

- People and organisations are accountable; tools are not.
- Oversight needs information, training and the authority to stop a system.
- The EU AI Act is risk-based, applies in stages, and requires organisations to support AI literacy; it sits alongside the GDPR.

Three practical takeaways

1. Know who in your organisation owns each AI tool you use.
2. Ask what AI-literacy support your organisation provides, and use it.
3. Report AI problems through your normal incident channels.

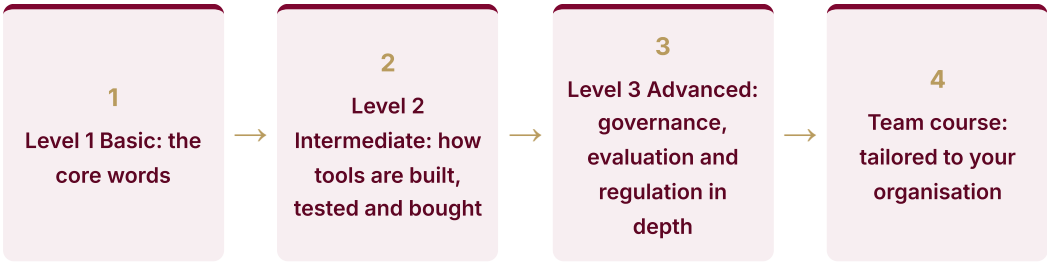
Self-check quiz, Chapter 8

1. Can "the AI decided" ever be an acceptable explanation for a decision?
2. What is shadow AI?
3. Since which date has the AI-literacy duty in Article 4 applied?
4. What is the difference between a provider and a deployer?
5. Does the EU AI Act replace the GDPR?

Next steps

You now hold the core vocabulary of AI. That alone changes how you read the news, sit in meetings and judge the tools you are offered. Here is where to go next.

FIGURE 16. YOUR LEARNING PATH



Each level builds on the one before; the team course applies the vocabulary to your own tools and workflows.

Level 2: Intermediate. Moves from words to judgement: how models are evaluated, what to ask vendors, how retrieval, fine-tuning and agents work in practice, and how to read a performance claim.

Level 3: Advanced. For those who govern, buy or build AI: risk management, validation design, monitoring, regulatory obligations and the questions boards should ask.

Both are available in the WyteRock AI Collection: https://wyterock.com/books.html

For your team. WyteRock designs customised AI-terminology and AI-literacy courses for clinical and non-clinical teams, using your own tools, policies and examples, and helping your organisation take proportionate measures under Article 4 of the EU AI Act.

- AI course for doctors and clinical teams: https://wyterock.com/ai-course-for-doctors.html
- Request a course or a specific task: https://wyterock.com/task-request.html

One habit to keep. Whenever you meet a new AI word, ask the four questions this book has used: What does it mean? What is it like? Where would I see it? What do people get wrong about it?

You are the pilot. CentiAgent is your co-pilot.

Answers to the self-check quizzes

Chapter 1. (1) AI. (2) False; they were loosely inspired by the brain, but the resemblance is mostly in the name. (3) The number of layers in the neural network. (4) No; it describes who is in charge, not how capable the tool is. (5) Traditional software follows rules people wrote; machine learning learns its rules from examples.

Chapter 2. (1) Training data is used to learn; test data is kept separate to check performance on unseen examples. (2) Using a trained model to produce an output. (3) It may have overfitted, memorising quirks rather than learning patterns that generalise. (4) The internal settings a model adjusts during training, which store what it has learned. (5) False; training usually happens beforehand, though product terms on using your inputs should be checked.

Chapter 3. (1) No; a token may be a word or part of a word. (2) Prompting changes the instruction; fine-tuning changes the model. (3) False; the context window is not memory. (4) It searches a trusted set of documents. (5) False; checking whether it is correct remains your job.

Chapter 4. (1) No; the tool has no intention, and hallucination is a known limitation. (2) Training data that under-represents certain groups (other answers, such as biased labels, are also correct). (3) Gradual decline in performance as the real world changes from the training data. (4) False; performance can drift, so monitoring is needed. (5) Tracing every important claim to its source.

Chapter 5. (1) No; many follow fixed scripts. (2) An assistant helps on request; an agent plans and carries out several steps towards a goal. (3) The human, the pilot. (4) Yes; much automation follows fixed rules. (5) When they approve without genuinely reviewing, or lack the time, information or authority to say no.

Chapter 6. (1) The tool says "no" when the condition is actually present: a missed case. (2) Its sensitivity is perfect, because it catches every real case, but it raises a false alarm for every healthy person, so its specificity is zero. (3) Usually sensitivity, because missing a case is serious. (4) A tool that always says "no" is right for most people when few have the condition, while missing every real case. (5) When too many alerts, many false, lead people to ignore them.

Chapter 7. (1) Yes. (2) Any two of: health data, genetic data, biometric data used for identification (others listed in Article 9 are also correct). (3) False; consent is one of six lawful bases. (4) Use only the personal data that is adequate, relevant and necessary for the purpose. (5) The right, with exceptions, not to be subject to decisions based solely on automated processing with legal or similarly significant effects.

Chapter 8. (1) No; people and organisations remain accountable. (2) Use of AI tools at work without the organisation's knowledge or approval. (3) 2 February 2025. (4) A provider develops or places an AI system on the market; a deployer uses it under its own authority professionally. (5) No; both apply.

Glossary

Accountability: a named person or organisation answers for how AI is used and for decisions made with it. **Accuracy:** how often a tool is right overall; can mislead when the condition is rare. **Agent (agentic AI):** an AI system that plans and carries out several steps towards a goal, often using other tools. **AI assistant:** a tool that helps with a range of tasks on request. **AI literacy:** the skills, knowledge and understanding to use AI sensibly, aware of its opportunities and risks; Article 4 of the EU AI Act requires measures to support it. **Alert fatigue:** ignoring alerts because too many are false alarms. **Anonymisation:** making data impossible to link to a person by any reasonably likely means. **Artificial intelligence (AI):** the broad field of making machines perform tasks that seem to need human intelligence. **Augmented intelligence:** AI that supports human judgement rather than replacing it. **Automated decision-making:** decisions made solely by automated means; GDPR Article 22 gives people rights about these. **Bias:** a systematic tendency to produce unfair or less accurate results for some groups. **Chatbot:** a program that holds a conversation with a person. **Co-pilot:** an AI tool built into your work that suggests and drafts while you decide. **Consent:** one of six lawful bases for using personal data under the GDPR. **Context window:** the amount of text, in tokens, a model can consider at once. **Data minimisation:** using only the personal data necessary for the purpose. **Dataset:** an organised collection of data gathered for a purpose. **Deep learning:** machine learning using many-layered neural networks. **Deployer:** under the EU AI Act, an organisation using an AI system under its own authority professionally. **Drift:** decline in a model's performance as the world changes. **DPIA (data protection impact assessment):** a risk assessment required by the GDPR for high-risk processing. **EU AI Act:** Regulation (EU) 2024/1689, the EU's risk-based law on AI, amended by Regulation (EU) 2026/1744. **False negative / false positive:** a missed case / a false alarm. **Fine-tuning:** further training of a model on specialised data. **GDPR:** the General Data Protection Regulation, Regulation (EU) 2016/679, applying since 25 May 2018. **Generative AI:** AI that creates new text, images, audio or code. **Hallucination:** a plausible but false or unsupported output. **Human in the loop:** a person reviews and approves AI outputs before they take effect. **Human oversight:** the ability of people to understand, monitor, intervene in and stop an AI system. **Inference:** using a trained model to produce an output. **Large language model (LLM):** a model trained on large amounts of text to predict likely next words. **Lawful basis:** the legal reason allowing personal data to be used. **Machine**

learning (ML): AI that learns patterns from examples. **Model:** the trained system that turns inputs into outputs. **Neural network:** layers of simple connected units that learn by adjusting their connections. **Output:** what an AI system produces; a suggestion, not a decision. **Overfitting:** learning training examples too closely, so new cases suffer. **Parameters:** the internal settings a model adjusts during training. **Personal data:** information relating to an identified or identifiable living person. **Prompt:** the instruction or question given to a generative AI tool. **Prompt engineering:** writing and refining prompts for better outputs. **Provider:** under the EU AI Act, whoever develops an AI system or places it on the market. **Pseudonymisation:** replacing identifiers with codes that can be re-linked with a key; still personal data. **Retrieval-augmented generation (RAG):** searching trusted documents first, then writing an answer from them. **Sensitivity:** how well a tool catches cases that are truly present. **Shadow AI:** AI used at work without approval. **Special category data:** more strictly protected personal data, including health, genetic and identifying biometric data. **Specificity:** how well a tool correctly clears cases that are truly absent. **Test data / training data:** data kept aside to check performance / data used to learn. **Token:** a small chunk of text a language model reads and writes.

Further reading and sources

Checked on 2 November 2026. Regulatory timelines change; recheck them before relying on them. All people, patients and organisations in examples and screen illustrations are fictional.

European Union regulation

- Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) (entry into force 1 August 2024; Articles 3, 4, 5, 6, 50 and 113; Annexes I and III). https://eur-lex.europa.eu/eli/reg/2024/1689/oj
- Regulation (EU) 2026/1744 amending Regulation (EU) 2024/1689 (Digital Omnibus on AI), published 24 July 2026, in force 27 July 2026; high-risk obligations from 2 December 2027 (Annex III) and 2 August 2028 (Annex I). https://eur-lex.europa.eu/eli/reg/2026/1744/oj
- European Commission, AI Act Service Desk. Article 4: AI literacy (consolidated text as amended). https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-4
- European Commission, AI Act Service Desk. Article 50: Transparency obligations for providers and deployers of certain AI systems. https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-50
- European Commission. AI Act: regulatory framework for AI (overview of the risk-based approach and timeline). https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
- Regulation (EU) 2016/679, General Data Protection Regulation (applying from 25 May 2018; Articles 4, 5, 6, 9, 22 and 35). https://eur-lex.europa.eu/eli/reg/2016/679/oj
- European Data Protection Board. Guidelines and recommendations on the GDPR. https://www.edpb.europa.eu/our-work-tools/general-guidance/guidelines-recommendations-best-practices_en

Concepts and measurement

- National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1 (January 2023), for definitions of trustworthiness, bias and oversight. https://doi.org/10.6028/NIST.AI.100-1

- National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, NIST AI 600-1 (26 July 2024), including "confabulation" (hallucination) as a generative-AI risk. https://doi.org/10.6028/NIST.AI.600-1
- Altman DG, Bland JM. Diagnostic tests 1: sensitivity and specificity. *BMJ* 1994;308:1552. https://doi.org/10.1136/bmj.308.6943.1552

About WyteRock

WyteRock Global is a France-based advisory firm, led by CEO Alex Mansour, helping organisations in health and life sciences adopt AI safely and practically.

We help clinicians, managers and their teams build AI literacy and use AI as augmented intelligence, with CentiAgent as co-pilot and people always in the pilot's seat. You are the pilot. CentiAgent is your co-pilot.

Members of the CentiAgent Club read every volume in the collection online: https://wyterock.com/centiagent-club.html

Learn more: https://wyterock.com/books.html · https://wyterock.com/ai-course-for-doctors.html · https://wyterock.com/copilot.html · https://wyterock.com/task-request.html