



Generative AI Principles & Guidelines

NB: it's hard to draw clear lines about what is and isn't AI but this document is focussed on *generative* AI – the tools that allow us to create content or interrogate complex questions.

This version: March 2026

How we approach Generative AI

We can do amazing things with AI. For starters, we can:

- Cut down on time-consuming, repetitive tasks and make time for more exciting, creative thinking.
- Make our content quicker, better and accessible to more people.
- Discover new ideas and approaches.

But we can also...

- Use content that we don't have the right to use or share content we should keep confidential.
- Reinforce stereotypes, cultural biases and false facts.
- Use-up the earth's resources for no good reason.
- Damage our understanding of our work by outsourcing our thinking to an algorithm. The more you devolve tasks to AI the less you are actively thinking about it and developing your own capabilities and opinions.

How do we make sure we do the good things?

We need to know how and why we are using Gen AI and that we are doing it with purpose and mindfulness, rather than it just happening.

The speed of change and size and nature of our organisation mean it's not practical to have comprehensive guidelines that cover all eventualities. But we need to use Gen AI, to do it in a thoughtful, conscious and ethical way; and to moderate, share and learn from what we're doing. To do this we have:

- This doc - our AI principles and practical guidelines doc: for external/ internal use.
- Short summary of how people in the organisation are using it – for internal/ external use – updated every 3/6 months.
- Internal survey every 6 months on what we're using and how, which is the basis for an online session for everyone to ask questions and share learnings – internal. This includes monitoring suppliers/ contractors' policies (eg: Smartsheet).
- Internal gen AI queries: raise with tech support via usual internal processes. If you're planning to use something new, this is where you should raise it.

Principles

Be informed about what you're using

- Know when you're using AI. Sometimes it's obvious, sometimes it might come as standard (like Co-pilot) or it might not feel like AI (like a picture editor or a note taker in a meeting).
- Understand that AI IS taking information and data from you and may not know itself how it will use that data. This doesn't just apply to uploaded docs but to also to prompts and to transcriptions/ meeting recordings.
- Ensure accessibility and inclusion: don't use AI to create resources/ options that don't work for everyone and avoid tools that have known biases (eg: Grok)

And think about why?

- What are you trying to achieve?
- What output do you want?
- How does it fit into your workflow?
- Is it the best tool for the job?

Be safe and legal

- Follow our data protection rules and copyright guidelines with all the tools you use.
- Only input necessary information and avoid sensitive or confidential data.
- Know what "sensitive" and "secure" means in practical terms (see guidelines below).
- Get your settings right. Even when these are right, don't share confidential information (see below).

Keep the human

- Be accountable. Always review your documents before using or sharing.
- Be representative and alert to bias in sources and sampling
- Check accuracy and bias: verify facts and review outputs to ensure fairness and quality.
- Create output that is fair and accessible to everyone.
- Don't assume the AI knows better (or writes better) than you!

Be transparent

- Let colleagues, partners, and service users know when AI is involved.
- You *don't* have to tell people that you got your first draft written in AI. You *do* have to tell them that you're using AI to transcribe a meeting they're attending

Practical Guidelines

How to Tell if You Might Be Using AI...

Sometimes it's obvious, sometimes it isn't – so keep an eye out. Things to watch out for:

- The tool produces text, images, summaries, or ideas automatically and quickly.
- The output may be generic, repetitive, or oddly confident.
- The platform mentions “machine learning”, “smart suggestions”, “auto-draft”, or similar terms. Other terms to watch out for: “assistant”, “co-pilot”, “generate”, “enhance”, or “magic” appear.
- Suggestions pop up automatically, such as predicted sentences or recommendations.
- There's no clear human author, or it's part of a known AI-enabled tool.
- A meeting is being recorded/ transcribed – openly by the meeting holders or it's being attended by notetakers (more on this below).

Before Using AI: Questions

- Do I really need AI for this task, and is it the right tool?
- Am I potentially entering personal, sensitive, or confidential information?
- How will I check the accuracy, fairness, and quality of the output?
- Could using AI here unintentionally exclude or disadvantage anyone?

Examples of how not to use AI

- **Making decisions without oversight:** using AI to decide who receives support, approves grants, or allocates resources without human review.
- **Sharing sensitive information:** inputting personal, confidential, or vulnerable individuals' data into AI tools that aren't secure.
- **Biased or harmful outputs:** publishing AI-generated content without checking it, which could unintentionally stereotype, misinform, or offend communities.
- **Mission conflict:** using AI to cut corners in ways that reduce service quality, mislead partners, or create inequities, even if it saves time.

Why is it dangerous to share confidential information with AI tools?

- You lose control of the information once it leaves your organisation, increasing the risk of accidental exposure or misuse.
- Your data will be stored or logged, even if not used for training.
- Cloud-based AI systems can be targeted by attackers, making sensitive data a security risk if ever compromised.
- Sharing confidential data could breach GDPR, NDAs, or internal policies.
- AI models may surface or reflect data elsewhere – there is a real risk of company data being shown to others.

- AI is an evolving tech and no-one, not even AI, really knows how this data might end up being used.

What information should not be shared?

NB: Sharing means uploading a file but also asking a question that includes any of this data or talking openly in meetings that are being transcribed/ attended by notetakers.

- **Personally Identifiable Information (PII):** Names, addresses, emails, phone numbers, dates of birth, passport numbers, National Insurance numbers, bank details.

Rule of thumb: If it identifies a real person, don't enter it.

- **Sensitive Personal Data:** Health information, religious or political beliefs, sexual orientation, biometric/genetic data, criminal records.

If it's private even to close friends, it's too sensitive for AI tools.

- **Confidential Work/Business Information:** Internal documents, product plans, financial forecasts, proprietary algorithms, internal emails, project details, meeting transcripts for summarisation.

NB: you cannot copy a transcript from one app (eg: Teams) into another (eg: ChatGPT) for a summary.

If sharing it externally would breach policy or NDA, don't put it in an AI chat.

- **Client, Customer, or Partner Information:** Names, financial info, legal cases, patient/student/tenant data, contracts.

Treat all client information as confidential by default.

- **Security-Sensitive Information:** Passwords, API keys, server/network details, vulnerabilities, incident reports.

If it relates to security, assume it's strictly confidential.

- **Information Covered by Legal or Contractual Obligations:** NDAs, regulated industry data, audit findings.

If you'd hesitate to email it externally, don't put it in AI.

How to make information secure

- Remove identifying details: replace names and specifics with placeholders (e.g., “Person A”, “Company X”).
- Summarise rather than copy-paste: describe the issue in general terms.
- Use approved enterprise versions: work accounts often offer enhanced data protection.

How to check your settings on ...

NB: Even with these tools properly set, do not share confidential information!

If you're starting with a new tool, please contact Kat/ Karen for a quick checklist on what to watch out for.

With tools you already use, find out how to control the settings by Googling:

"change data sharing settings in <tool name>"

And amend accordingly!

Remember that Google will probably give you an AI response which may or may not be accurate. So, better to look down the search results to find the help content pages for the tool you are using.

NB: At DU, we all have a basic version of MS365 which covers Teams, Mail and a web version of Excel/ Word. The settings on these are controlled by the main admin.

Most of us also have a personal version of MS365 – which may have Co-pilot on it depending on your subscription – which are settings you need to control yourself.

AI notetakers

AI notetakers are becoming increasingly common in online meetings. These can be built-in tools which aren't always detectable to other attendees (like Copilot in Teams), or third-party AI notetakers (such as Otter or Fireflies), which attend the meeting as a participant.

They can be really helpful, especially for people with accessibility needs, but they can also record people's names and/or images without their consent, and the data can be stored or shared without their knowledge. This is unethical and illegal.

If you are attending a meeting that is being recorded/ transcribed/ attended by AI notetakers, think carefully about what information you share and how.

If you are running a meeting for DU, as a rule:

- We don't allow external AI notetakers in our meetings.
- If we are recording or transcribing the meeting, we ask for consent at the start; make clear with whom the data will be shared and give people the option to switch off their cameras/ change their names.
- If people have accessibility needs that mean they require a transcription, we ask them to contact us beforehand and we'll make sure one is captured with consent.
- We encourage people to have their cameras on, especially at the start. This is good practice anyway but also lets us make sure all attendees are real people.

There is a separate practical document with housekeeping scripts/ notes on how to change participants' settings in the policies folder.

Helpful resources

This is a helpful guide to using AI ethically and responsibly - it's focused on academic studies but is useful for everyone.

[Using AI ethically & responsibly - Using Generative AI Tools in Academic Work - Subject guides at The University of Edinburgh](#)

This is our own [DCN skills boost to recognising your own unconscious biases](#)

And this from the Government also gives a good overview on safety/security of AI [Guidance to civil servants on use of generative AI - GOV.UK](#)

How we're using AI at DU

As of Jan 2026 (updated every 6 months)

Like many organisations, how we use AI varies depending on the individual and role and that includes whether the individual is working for as an employee or with us as a contractor. Our guidelines cover both.

Although we are interested in how it can help us, like many people, we've found the quality of what it produces variable, and often factually unreliable – so it often doesn't help us as much as we think it will.

How we're using it

- Background research & first drafts for our general business

Most people at DU use ChatGPT, Gemini and/ or Co-pilot to write drafts and rewrites (for example, in more accessible language by lowering the reading age), summarise information and research, generate ideas (“10 icebreakers for a community get-together”) and outline initial project plans. Everything is then reviewed by a human and edited/ adapted so that all our outputs rely on and reflect our personal knowledge, context and understanding, and individual craft and style.

- Admin

We don't use much AI in our admin. All interactions with DU are with people.

- Learning content

Our courses are written by humans. We use gen AI for specific elements of our learning content: we use Vyond to create short animations to sit within the courses, and we sometimes use AI voice-overs. We have used Rise 360 to help create/ re-write courses but are not currently using it.

- Generating code/ digital development

We use it to generate code for our sites/ platforms. A human then reviews, tests and incorporates it before deploying it. We ensure our suppliers/ the platforms are compliant with our standards.

- Marketing/ promotional material

We don't use it to generate text (as it's usually very generic and not very good). We sometimes use it to create images, particularly in Canva, but we've found it variable and not that helpful to use. We have used AI to generate voice-overs for our promotional films.

- Analysing Data

We don't use it in reporting/ analytics for our own platforms/ activities, but we do use it to analyse publicly available data to inform projects. For example, we've used it analyse data to help councils identify areas that would benefit from Digital Champion projects.

- Suppliers/ contractors/ online tools

We check the AI policies of our suppliers to ensure they are aligned with our own. This is part of our regular 3-monthly review process.

How this document was created

We used AI to generate a first draft of this document, but relied more heavily on these useful resources to inform it:

AI policy template via CAST: [AI-Policy-Template-by-ANB-Advisory.pdf - Google Drive](#)

Simple principles statement from the National Lottery Community Fund [AI principles](#)

And a huge raft of resources here [AI resources](#) from CAST.

SCVO [AI organisational policies - SCVO](#) and [SCVO-AI-guidelines-v1.0.pdf](#)