INSTITUTE FOR
Ethics in AI

Digital
Care Hub

Casson Consulting
Encouraging innovation in social care

The responsible use of Generative AI in adult social care: A value-led approach

Executive Summary

"The 'responsible use of (generative) AI in social care' means that the use of AI systems in the care or related to the care of people supports and does not undermine, harm or unfairly breach fundamental values of care, including human rights, independence, choice and control, dignity, equality and wellbeing." (The Oxford collaboration's value-led definition of the responsible use of (generative) AI in adult social care)

Adult social care supports individuals with disabilities, illnesses, or other needs, enabling them to live independently with dignity, choice, and respect for human rights. This includes both practical assistance, such as personal care and food preparation, as well as emotional and social support. Care services are delivered by formal care providers in care homes, day centers, and individuals' homes, as well as by informal caregivers, including family and friends. Central and local governments play crucial roles in shaping policy, providing financial support, and regulating care quality. Artificial Intelligence (AI) encompasses technologies that replicate human cognitive functions such as learning and decision-making. Generative AI creates new content—including text, images, and videos—based on prompts. Popular general-purpose AI tools like ChatGPT and Microsoft Co-Pilot, as well as social care-specific AI solutions, are transforming how care is planned and delivered.

Current use cases of generative AI in social care include:

- Assisting in generating care plans, meeting notes, and activity schedules.
- Supporting non-native English speakers in written communication.
- Checking health symptoms for preliminary insights.
- Managing administrative tasks like emails and letters.
- Providing AI-powered chatbot support for mental health and well-being.

While generative AI offers potential benefits to people in social care, it also poses many ethical considerations and risks due to technical limitations as well as related to how people may use generative AI in social care.

The Oxford collaboration on the responsible use of generative AI in adult social care launched in February 2024 at the University of Oxford, with the aim to define what the responsible use of generative AI in adult social care means and to address the gaps in official guidance and support. The collaboration included over 70 individuals and organisations from across the care community, including people who draw on care and support, careworkers, care providers, tech developers, academics, advocacy groups and policy makers.

We co-produced a value-led approach to the responsible use of (generative) AI in social care, with a focus on what care can do for people and on fundamental values of care. We published guidance and a call for action to define future steps. All resources are available here:

https://www.digitalcarehub.co.uk/ai-and-robotics/oxford-project-the-responsible-use-of-generative-ai-in-social-care/

1. Introduction

1.1 Background to the project

Social care

Adult social care is about supporting people with disability, illness or other needs to live as independently as possible, with choice and control, in dignity and with their human rights respected. Social care encompasses both practical assistance with activities of daily living, such as personal care or food preparation, and social and emotional support. Social care is offered by paid formal care providers in settings such as care homes or day centres as well as people's own homes or by unpaid caregivers, like family members or friends. Central and local government play an important role in adult social care provision, through their policy directions, financial support for individuals in need of care and support and carers and as regulators of the quality of paid care provision.

AI and 'generative' AI

'Artificial Intelligence' does not have one agreed upon definition. However, it can loosely be understood as a term that refers to machines that mimic human cognitive functions such as learning, problem-solving, and decision-making. There are different types of AI, including computer vision or machine learning. Popular examples of AI systems are systems that make driverless cars possible, or systems that can identify potential cancerous growths from body scans.

AI is not new, with decades of development to look back on. But powerful new 'generative AI 'systems that have emerged in the past years have been revolutionary in terms of accessibility, usability and broad impact, with many people changing how they go about their work and tasks of daily living because of these new systems. 'Generative AI' refers to a type of AI under the category of 'machine learning'. Generative AI creates new text, videos, pictures and more when prompted to do so. Prominent examples of general-purpose generative AI systems on the market are OpenAI's ChatGPT or Microsoft's Co-Pilot. These systems create new texts when asked to do so by a user, which mimic human language through powerful word prediction. These word prediction systems are also referred to as "Large Language Models" (LLMs).

Next to general purpose-systems such as ChatGPT, there are also generative AI systems specifically created for social care purposes, for example to create care notes or activity plans, to help people express themselves and to act as 'companions' .

In this collaboration we have focused on generative AI because of its wide availability, many potential use cases across social care and many new generative AI products coming onto the market for social care. However, the definition of 'responsible use of AI' in this white paper and many of the documents and resources we created under this collaboration equally apply to other types of AI in social care. Still, we continue to focus on generative AI in this section of the white paper as a case study of a widely available AI system, that can transform people's care.

Technical limitations of generative AI

Generative AI is improving constantly, with increasing capacity to produce impressive responses to all sorts of prompts, including care-relevant ones. Generative AI may therefore support practical aspects of care, such as planning activities and writing up notes. But the use of generative AI in the provision of care raise some important ethical questions. To understand the ethical considerations relating to the use of generative AI in care, it is important to shine a light on some of the *technical limitations* of generative AI, as well as potential issues regarding *the way people use* generative AI in care:

Examples of relevant technical limitations/considerations include:

- **Understanding and contextualisation**

LLMs, as convincing and human-like as outputs may sound, are based on algorithms that are very good at predicting word sequences. But these systems do not have an 'understanding' of the world around us and they cannot contextualise. They are not 'human' intelligent. This means that when used in social care, these systems may produce solutions or ideas that sound reasonable, but when put in context, they don't make sense or in the worst-case scenario could be harmful or present a safeguarding issue.

- **'Hallucinations'**

Another limitation of generative AI is the phenomenon of "hallucinations," where the model produces outputs that may appear to us as true and/or reasonable, but in fact are made up and untrue. These hallucinations can range from minor inaccuracies, such as incorrect historical facts, to severe and legally actionable fabrications, such as dangerous medical advice (although many LLMs in the public domain are deliberately configured not to give medical advice).

- **Output Quality**

While outputs may sound reasonable, there is no guarantee that the information produced is indeed of high quality or reflects good practice. For example, a person prompting a LLM to offer some care-related solutions (e.g. How to help someone get up after a fall), the response may sound like a reasonable suggestion, but it may not indeed be the safest solution or one which follows good practice overall. Also, good practice changes in social care. Outputs of a generative AI model trained before the changes occurred may take time to catch up and, until then, may perpetuate practices that are no longer advocated.

- **Data Privacy**

Developers and deployers of generative AI are subject to Data Protection laws. However, generative AI systems may have been both intentionally and unintentionally trained on personal information, which could be mishandled or exposed if not properly protected. Even anonymised data can sometimes be re-identified through sophisticated analysis. The inner workings of many generative AI models are complex and opaque. This can make it difficult for users to understand how their data is being collected, used, stored, accessed or shared, leading to potential privacy concerns. Sometimes, AI models might inadvertently generate outputs that include or infer personal data, especially if the models were trained on data containing such information. This can lead to unintentional privacy breaches.

- **Biases**

Generative AI is susceptible to biased outputs due to biases present in the training data[1]. This can perpetuate and amplify existing prejudices, which pose a risk to people drawing on care and support as well as posing ethical and reputational risks for social care organisations. The problem of biased outputs becomes even more problematic in scenarios, in which the person who has prompted the system and is acting on the output is biased themselves. Existing inequalities could therefore be quickly reinforced.

Use cases of generative AI in social care

As part of this collaboration, we have interviewed care workers, working directly with people who draw on care, and listened to care providers, people who draw on care and support and family caregivers to understand better how they use AI - particularly generative AI - in caregiving. Many care workers told us that they have not yet started to use generative AI in their work, mainly because they are unsure about policies and because they have other systems in place to support them with their work. However, we also often heard of following use-cases:

- Using general-purpose generative AI (e.g. ChatGPT or similar) or social care specific products to generate activity plans, care plans or meeting notes in residential or domiciliary care contexts. We heard that this use is particularly prevalent among care workers, who do not have English as their first language, to support written tasks.
- Using generative AI to check health symptoms of people drawing on care and support to get ideas about conditions they may be suffering from.
- Using generative AI to help with administrative tasks, like letter or email writing.
- Using generative AI powered chatbots for the purposes of mental health support e.g. talking to chatbots about particular problems people were facing.

People told us how these specific use cases helped them to deal with heavy administrative loads, empower them to speak up or find new solutions and insights. However, many were also concerned that they didn't fully understand technical limitations or were worried about breaching data privacy.

The Care Workers' statement, produced by our care workers' working group, outlines some of the needs and demands of care workers for them to be able to use generative AI confidently and safely in their work.

Policy and regulation of AI in social care (state: April 2025)

The UK government has embraced the opportunities that AI may bring and is positioning itself as a leader in AI innovation. In January 2025, it announced its AI Action Plan[2] and with it the intention to roll out AI systems across public services. Some local authorities have started to integrate AI systems

[1] See for example an output from Google's Gemini AI: https://www.cbsnews.com/news/google-ai-chatbot-threatening-message-human-please-die/
[2] AI Opportunities Action Plan https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan

into their work, for example through offering chatbots to people to answer questions around parking fines, rubbish removal[3] etc.

From a regulatory point of view, there is currently no 'AI law' in the UK (although UK based business may be subject to the EU's AI Act). Rather, various regulators like the Information Commissioner[4] or the Competition and Markets Authority (CMA)[5] have published their own regulatory directions into AI and generative AI more specifically. However, regulators of health and social care in the UK, including the Care Quality Commission (CQC) have not yet published any clear guidance on how to integrate AI into social care provision and the implications of AI for the current legal and regulatory system of social care (e.g. Care Act 2014, mental capacity law) is yet to be determined.

There is therefore currently a lack of official guidance on how the interaction between current regulation of social care and AI. There have been various initiatives to address the need for good practice understanding and guidance across social care, including this one. However, in order to make the most of what AI and generative AI has to offer to people who draw on care and support as well as people who provide care, it is important to provide clarity around the meaning of the responsible use of AI in social care contexts.

1.2 The Oxford collaboration on the responsible use of generative AI in adult social care

The 'Oxford collaboration on the responsible use of (generative) AI' ('The Oxford collaboration) refers to a cross-care community collaboration that began in February 2024 following a roundtable event held at the University of Oxford to discuss the implications of generative AI on care and the publication of a shared statement This statement was endorsed by over 30 organisations. The aim of the collaboration has been:

- to co-produce an understanding of what the 'responsible use of (generative) AI' in adult social care means and;
- to address a lack of official guidance and support for people with lived experience, care providers and frontline care workers in this area.

The collaboration grew beyond the initial signatories of the statement and included more than 70 organisations and people who draw on care and support , family caregivers, care providers (domiciliary/residential care/supported living), care workers, tech suppliers, academics, policy makers in local and national government, and wider civil society mainly from England, but also Scotland and Wales. Social workers were also part of this collaboration.

Governance structure

The collaboration was governed by a steering committee with representatives from across the care community and working group leads, who volunteered to move forward five working groups (see

[3] See this website for more information
[4] Information Commissioner's Office https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/
[5] Competitions Market Authority https://www.gov.uk/government/publications/cma-ai-strategic-update

table below for an overview of the collaborations' governance and structure). The working groups enabled individuals with shared professional backgrounds to convene, explore the meaning of generative AI for their work and lives and identify the outputs that they would like to produce. The outputs of all working groups were then subject to an overarching co-production working group, that scrutinised and deliberated on the overall outcomes of the collaboration. We also convened a deliberation for the entire collaboration at the University of Oxford, to decide on our definition of 'the responsible use of (generative) AI in social care' (see page 9 below), our guidance and call for action.

Table 1: Governance structure of the Oxford collaboration on the responsible use of generative AI in adult social care

Overarching	**Outputs**
Hosting	Overall administration
Steering group	Strategic direction
Working groups	
People with lived experience	Co-produced overall guidance
Community	Statement
Care workers	Statement
Care providers	Care providers' guidance
Tech suppliers	Tech suppliers' pledge
Principles	Foundation principles
Evaluation framework	Evaluation of AI governance in formal care (not yet published)

At the heart of the Oxford collaboration were following core principles:

- **Co-production:** 'Co-production' is a specific approach to responding to a particular problem or issue, in which people with lived experience are equal contributors alongside care workers and others. The definition of the 'responsible use of (generative) AI' must be shaped together by people with lived experience, care workers as well as care providers and tech suppliers, policy makers and others in the care community in order to find solutions and practices that make sense for everyone. Processes of co-production are therefore inclusive, in which people with different experiences and backgrounds cannot only make 'their voices heard', but actively shape discussions and outcomes.

- **Collaboration:** AI is developing rapidly and being rolled out across social care, with an urgent need to understand the potential benefits as well as harms and how to mitigate risks. We believe that the care community (including tech suppliers and policy makers) collaborating with each other will ensure that we can respond most effectively to the need for clarity and guidance, avoiding fragmentation of good practice understanding and thereby building confidence in people who draw on care and support, unpaid carers, care workers, care providers and other key actors.

- **Transparency, inclusivity and ethics:** Across our collaborative activities, we maintained an inclusive approach, meaning that any organisation or individual could join the process including the steering group if the individual/organisation. We did maintain special terms of

participation in the steering group for private tech companies in order to avoid unfair competitive advantage. Documents and information about the collaboration were published and updated on a regular basis and communicated at various care sector specific events and fairs. We kept publishing information about the project and working documents on the Digital Care Hub's website. An overall risk assessment was undertaken for this collaboration and specific activities, including the interviews with careworkers, were subject to ethical clearance by the University of Oxford's Research Ethics Committee.

---

2. The meaning of 'responsible use' of (generative) AI in social care

'Responsible AI'

The concept of 'responsible AI' is not uniquely linked to this Oxford collaboration, but generally an important concept in AI ethics. It refers to a particular understanding of AI as technology that automates tasks that will support people, but still requires human inputs to develop and function. AI does not exist in a silo, but in a complex societal system of power structures and human decisions on who develops AI, how it is used, who has access to it etc. 'Responsible AI' is therefore about being aware and reflective of possible and actual impacts related to AI [6] on people and society.

From this perspective, 'responsible use of AI' requires actors to take responsibility for the impacts that can arise from AI. It also needs systems to hold people and organisations accountable. This responsibility and accountability does not lie with one group of people, like tech developers, alone. Rather, it is a shared responsibility across government, tech developers, users of AI etc.

The Oxford collaboration's definition of the responsible use of (generative) AI in adult social care

"The 'responsible use of (generative) AI in social care' means that the use of AI systems in the care or related to the care of people supports and does not undermine, harm or breach fundamental values of care, including human rights, independence, choice and control, dignity, equality and wellbeing."

When defining the responsible use of generative AI in social care, we started with a basic shared understanding of social care independent from AI or technology. We reminded ourselves what social care can offer people, centred around fundamental values and principles of care. We then asked ourselves the question how AI might be able to support caregiving and how AI - and generative AI specifically - could pose a risk to what we value in care. Our principles working group created a principles framework, which are reflected in our guidance.

We therefore took a 'value-led approach', rather than one that focuses on what problems AI might be able to solve in social care (e.g. staff shortages) or what outcomes we might wish to see from using AI (e.g. more effective working processes[7]). This approach does still allow for the identification

---

[6] See Victoria Dignum (2023). Responsible Artificial Intelligence: Recommendations and Lessons Learned. In: Eke, D.O., Wakunuma, K., Akintoye, S. (eds) Responsible AI in Africa. Social and Cultural Studies of Robots and AI. Palgrave Macmillan, Cham.

[7] Whitfield, Grace and Wright, James and Hamblin, Kate, AI in Care: A Solution to the 'Care Crisis' in England? (January 8, 2024). Forthcoming, Edward Elgar Handbook on AI and Public Policy, edited by Regine Paul, Emma

of specific problems and desired outcomes of the use of (generative) AI in social care, but it puts these in a wider ethical context and value-based understanding to determine benefits and risks.

We recognise that the definition gives rise to responsibilities on the side of various actors, including tech developers and companies offering AI solutions, to policy makers and central/local government and to care providers. We will require adapted and possibly new systems to hold people and organisations accountable, with the need for a good understanding how current regulation of social care and beyond is 'fit for purpose'.

3. Future steps

AI is rapidly developing and uptake of AI systems across social care is increasing. We recognise that this work was only the beginning. **We have therefore announced a new Alliance on the responsible use of AI in social care.** We also call for following action:

We are a group of people who draw on care and support, people who provide unpaid care for friends or family members, care workers, care providers, strategic leaders in social care, providers and developers of technology **and others who have spent time individually, and collectively, considering how Generative AI can be used responsibly and ethically in social care.**

The 'responsible use of (generative) AI in social care' means that the use of AI systems in care or related to the care of people supports and does not undermine, harm or unfairly breach fundamental values of care, including human rights, independence, choice and control, dignity, equality and wellbeing.

We are concerned that Generative AI is being adopted at pace, without clear guidelines or guardrails to ensure that its use is responsible, safe and effective. This rollout is often a response to pressures on the current system of care and support, and this may detract from the focus on human rights, equality and legal frameworks which ought to inform its use.

We have developed our own guidance to start to address these gaps, but we also need the following actions to be taken:

1. Everyone – to use the "I" and "we" statements set out in our guidance, to guide the use of Generative AI in social care.

We're asking everyone with a stake in the use of Generative AI in social care to consider how you can apply the guidance ("I" and "we" statements) we have collectively produced, in your work. If you're a person who draws on care and support or someone who works in social care, you can use these "I" statements to check whether technology is being used responsibly. If you're a provider of care (in a local authority, a care organisation or somewhere else), or a technology provider or developer, you should consider whether the "we" statements reflect how you'd describe your work.

Carmel and Jennifer Cobbe, 2024, Edward Elgar Publishing Ltd., Available at SSRN: https://ssrn.com/abstract=4687695 or http://dx.doi.org/10.2139/ssrn.4687695

2. Everyone working on Generative AI in social care – to continue to work collaboratively on key issues.

We've learned a lot by bringing together different perspectives across social care, and we know we're not the only ones thinking about this and work is ongoing across adult social care, in relation to children and young people and in other related fields. We're keen to continue the collaboration, making sure that we're bringing different perspectives together, and crucially ensuring that there is active participation by people who draw on care and support, unpaid carers, people who work in social care and others, as we move forward. We know there is still significant work to do, including to address issues around cybersecurity and the environmental impact of Generative AI.

3. UK governments to work with the UK's regulators to develop appropriate regulatory and accountability structures to govern the use of Generative AI (and other AI technologies) in social care.

Our discussions have laid bare the need for clear and enforceable guidelines around the use of Generative AI in social care. We also need clearer mechanisms for accountability. We envisage a role for UK care regulators, local authorities in relation to their safeguarding responsibilities, and professional bodies linked to the social care sector. UK governments need to recognise the current regulatory gaps and designate a responsible body to address these. We believe our guidance can provide a helpful starting point for this work.

4. UK governments to take a lead in developing and maturing the infrastructure for innovation and entrepreneurship in social care, in partnership with other key stakeholders including local authorities.

We need to create a supportive ecosystem and infrastructure which can enable high quality, inclusive innovation and entrepreneurship in social care technology. We need to build infrastructure to support inclusive, human-centred design and to enable and support coproduction and involvement across the innovation life cycle.

5. UK governments to take a lead in developing and nurturing new business models in the social care technology field, in partnership with other key bodies including local authorities.

The fragmentation of the social care sector makes it hard to drive economies of scale in care technology development. Power is not held equally between large, often national or global, technology providers; small, often local, care providers; and care workers and people who draw on care and support. Individuals, communities and local economies do not always feel able to share in the rewards from technologies built using their data, and around their lives. We believe governments can play a role in supporting the development of new business models in the care technology field, which are more efficient better support growth across local communities and economies.

6. The Department of Health and Social Care to ensure that it's promised National Standards around the use of technology in social care are ethically informed and aligned with existing legal frameworks, including human rights and equality law, and the wellbeing principle established in the Care Act 2014.

The DHSC has promised new National Standards for the use of care technology in England. We hope that the work we've already done can offer a starting point for these in England and in the development of similar standards across the UK.

To access all the documents produced under this collaboration, please visit:

https://www.digitalcarehub.co.uk/ai-and-robotics/oxford-project-the-responsible-use-of-generative-ai-in-social-care/

Acknowledgements

This collaboration has included many individuals and organisations across the care community. We would like to thank all of them. A special thank you to the leads of our working groups: Dr Donald Macaskill, Dr Jane Townson OBE, Tyler Reinmund, John Boyle and Karolina Gerlich.

This white paper was authored by Dr Caroline Green, Katie Thorn, John Boyle, Kate Jopling and Daniel Casson. To cite, please use: Green, C., Thorn, K., Boyle, J., Jopling, K. and Casson, D. (2025) White paper of the Oxford Collaboration on the responsible use of generative AI in social care: A valued led approach.