

OPEN

Corporate Policy Committee

20 March 2025

Generative AI Policy Briefing

Report of: Adele Taylor, Interim Executive Director of Resources and S151 Officer

Report Reference No: CPC/61/24-25

Ward(s) Affected: All

For Decision or Scrutiny: Decision

Purpose of Report

- 1 This briefing report provides an update on the use (and potential use) of Generative Artificial Intelligence ("GenAI") by the Council and highlights issues, opportunities, and ethics to be considered in a local government setting and in support of the Corporate Plan for an effective and enabling council.
- 2 A draft GenAl Policy is attached with reference to the supporting governance framework particularly as this underlying policy will be in a constant state of development whilst the technology matures.
- 3 The Corporate Policy Committee is asked to approve the draft corporate GenAI Policy, within the context of the background information within this briefing report.

Executive Summary

4 This briefing report provides an update on the use (and potential use) of Artificial Intelligence by the Council, its employees, contractors, developers, vendors, temporary staff, consultants or other third parties, highlighting any issues, opportunities, and ethics to be considered in a local government setting.

- 5 There are three main uses of Artificial Intelligence (AI): Machine Learning, Predictive Analytics, and Generative. In the first instance the Council is focussing on Generative AI.
- 6 Generative AI ("GenAI") services such as ChatGPT for text, and others for creation of images, audio, and video present a risk if used without understanding the capabilities.
- 7 Given that GenAl is the most readily available for use by council staff in undertaking their daily work, this briefing asks that the Council recognises the need to approve policy within a framework of existing governance to empower the design and use of GenAl, and that it is ethical, complies with all applicable laws, regulations, and other Council policies, and compliments the Council's existing information and security policies.
- 8 The Information Assurance and Data Management (IADM) Programme and the Strategic Information Governance Group (SIGG) are existing governance frameworks which have developed data governance strategies and policies. The IADM operational programme has provided guidance and assurance for the adoption of GenAI across the authority. And the SIGG established a Task and Finish group to develop the GenAI Policy within this report.
- 9 It should be noted that the pace of development and application of GenAI is such that any policy will be in a constant state of development.

RECOMMENDATIONS

The Corporate Policy Committee is recommended to:

- 1. Approve the Corporate GenAl Policy and underpinning governance and related usage guidance, within the context of the background information within this briefing report.
- 2. Note that SIGG has corporate responsibility for data governance and that it will implement mechanisms for effective GenAI data strategies, policies, and procedures to ensure data accuracy, ethical use and individual privacy whilst ensuring adequate data collection and the means to regulate its use.
- 3. Note that the Digital Workstream is responsible for maintaining the GenAl Policy and that risks and issues will be managed as the technology matures.
- 4. Note that IADM continues to manage guidance and assurance to successfully adopt GenAI across the authority and will work with the Digital Delivery Programme to ensure that both AI technical solutions and the underlying information are compliant and secure.

Background

- 10 There are three main uses of Artificial Intelligence (AI):
 - **Machine Learning**: Classification, matching, risk assessments. Examples include detection of anomalies, fraud, diagnosis, voice recognition.
 - **Predictive Analytics**: Prediction of flow, modelling complex systems, simulations. Examples include protein folding, drug development.
 - **Generative**: Creates outputs that look original, may be text, image, voice, or video. Examples include creating summaries of documents, responding the prompts, generating images.
- 11 There is a lot of information on Generative Artificial Intelligence (GenAl) or Artificial Intelligence (AI) in the media, the Council therefore needs to understand the detail of what AI can and will do for the authority. Briefing paper will focus on GenAI as this tool is now readily available on the internet for staff to use and therefore its capabilities, limitations and when it should and should not be used need to be better understood.
- 12 Generative AI ("GenAI") services such as ChatGPT, Bard and CoPilot, and the likes of DALL-E, Midjourney and Stable Diffusion for image generation been in development for many years but have shot to prominence recently as they have become more publicly available. There is a risk that without understanding the capabilities and more importantly, the limitations of these services they could be used while undertaking Council business without the necessary checks and balances in place. In addition, the likes of DeepSeek and QWEN developed in China have led to security and governance concerns.
- 13 A secure, managed implementation of GenAl is available as part of the Council's existing Microsoft 365 investment in the form of CoPilot.
- 14 Al can broadly be divided into two categories, Narrow Intelligence, and General Intelligence. Narrow Intelligence is Al focussed on performing one main task. All current Al has Narrow Intelligence although they may appear smarter than they are, for example Alexa, which is able to perform a limited pre-defined set of operations. General Intelligence refers to machines that can perform many tasks, be cognitively aware of what they are doing, and are able to self-learn and adapt. Artificial General Intelligence ("AGI") is a long-term aspiration and is not readily available now. There are suggestions that AGI machines will replace human intelligence, but machine intelligence may be different <u>The Dstl Biscuit Book WEB.pdf (publishing.service.gov.uk)</u>

- 15 Further references to GenAI or AI in this briefing document will refer to Narrow Intelligence.
- 16 Al is commonly understood to be technological systems with the ability to process data in a way which resembles intelligent behaviour. This broad definition is important, as the rapid pace of technological change would quickly render any fixed, narrow definition outdated and make future-proof policies infeasible.
- 17 Agentic AI refers to AI systems that not only understand and respond to their environment but also possess the ability to autonomously, making decisions and adapting in real time without constant human guidance. Decision.
- 18 Al is driving transformation across all areas of society today. It is an umbrella term encompassing a range of technologies both sophisticated and simple that are used to, among other things, to make predictions, inferences, recommendations, or decisions with data. Al is used in many products and services that people use, interact with, or are impacted by, every day.
- 19 Al is already being used, for example, to help diagnose medical conditions and care for patients, to predict demand for transportation services, and to optimise energy consumption of devices and large systems. It can also perform energy management, to detect fraudulent financial transactions and manage financial services, and to recommend products or content and operate consumer internet and e-commerce platforms. It is predicted that Al will touch virtually every area of life in the years ahead, in ways both visible and invisible, and its impacts will be felt from a personal to a societal scale.
- 20 It is already an unstoppable force in our economy. <u>According to Tech Nation, there are more than 1,300 AI companies in the United Kingdom.</u> OpenAI, the developer of ChatGPT, has announced its first international office will be based in the UK. Research commissioned by the government and published in January 2022 shows <u>UK businesses</u> spent around £63bn on AI technology and AI-related labour in 2020 alone. This figure is expected to reach more than £200bn by 2040, when it is predicted more than 1.3m UK businesses will be using AI. The council will ensure that it is a position to take advantage of AI powered tools where appropriate to improve the way it delivers services.
- 21 In its <u>National AI Strategy</u>, the government has committed to developing a pro-innovation national position on governing and regulating AI and outlined their position in a <u>policy paper</u> in 2022. This white paper sets

out the government's proposals for implementing a proportionate, future-proof, and pro-innovation framework for regulating AI.

- 22 The <u>UK Science and Technology Framework</u> also sets out the government's strategic vision to make the UK a science and technology superpower by 2030. It identifies AI as one of five critical technologies and notes that regulation plays a key role in creating an environment for AI to flourish.
- 23 Through this innovative approach to AI regulation, central government is hoping to help the UK harness the opportunities and benefits that AI technologies present. The expectation is that this will drive growth and prosperity by boosting innovation and investment and building public trust in AI. And it will strengthen the UK's position as a global leader in AI, by ensuring the UK is the best place to develop and use AI technologies.
- 24 It is recognised that local government needs to transform digitally to meet the expectations of residents. A combination of digitally enabled customer services, underpinned by AI technologies, will enable public sector transformation.
- 25 It should be noted that some AI researchers argue that AI currently is just "statistics on steroids." All agree that AI is transforming how we work, live, wage war and even understand what it means to be human, as <u>Professor Stuart Russell</u> explored in his <u>BBC Reith Lectures</u> in December 2022.
- 26 Simplistically, GenAl Large Language Models (LLMs) are developed and train (or learn) from substantial amounts of text-based data, typically scraped from the open internet, which covers web pages and other sources such as scientific research, books, or social media posts.
- 27 The LLM (Large Language Models) then analyses the relationships between different words and turns that into a probability model. The LLM, through a tool such as a chatbot (ChatGPT is one such highprofile chatbot), will then generate text that appears the same as that generated by a human.
- 28 There are several types of chatbot that are used for different purposes:
 - (a) Retrieval chatbots: These chatbots provide an interface to structure content in the knowledgebase. Before implementing a chatbot, the knowledgebase must be mature and up to date.
 - (b) Generative chatbots: These chatbots submit, categorise, and route requests on the end user's behalf. These chatbots can be used to verify identity and gather requirements to save time for

live agents. This cuts down on time spent by live agents on every ticket.

- (c) Intelligent agents or virtual assistants: These are powered by machine learning to for example automate IT (Information Technology) operations like incident, problem, and change management and to identify issues early and resolve them before they have an impact. Both the live agent and the customer can also use intelligent agents to perform repetitive tasks needed for tickets.
- 29 ChatGPT is an example of a generative chatbot founded by the company OpenAI which has links to Microsoft and Amazon. Microsoft is committed to adding ChatGPT to Microsoft 365, it called CoPilot and be embedded into Word, Excel, PowerPoint etc.
- 30 ChatGPT when given the prompt "what an organisation should consider when implementing AI?" responded that "*The use of AI tools must adhere to principles of transparency and fairness. Companies should have policies in place to address data privacy and security, bias and discrimination, and transparency and explainability. Employees must use AI tools in a manner consistent with security best practice.*"
- 31 In summary, the Council needs to recognise:
 - (a) All is the single most exciting and potentially disruptive technology to emerge for a long time.
 - (b) It has been around for years, we have seen it coming, however there is a change in thinking underway – this might be the first time in human evolution where humans may be subject to the decisions of an entity that is not human and (unlike a standard algorithm) whose reasons for the decision are not fully explainable.
 - (c) Machine learning LLMs bring a whole new meaning to Isaac Newton's "standing on the shoulders of giants."
 - (d) The growth in AI deployment and integration is already well underway.
 - (e) Once this snowball starts, (it has) the momentum to accelerate is exponential.
 - (f) Are we ready for it?

Briefing Information

32 There has been consultation and engagement with the services and enabler representatives on the Digital Workstream, SIGG and IADM, in order that that this briefing paper can recommend the draft GenAI Policy and underpinning governance to the Corporate Policy Committeee (Appendix 1).

- 33 SIGG, the Task and Finish Group, and IADM have referred to best practice AI policies such as the one produced (in collaboration with <u>ALGIM</u>, Socitm's partner association in New Zealand) which provides a framework for the use of Generative Artificial Intelligence (GenAI) large language models such as ChatGPT, Bard, Bing or other similar tools by council employees, contractors, developers, vendors, temporary staff, consultants or other third parties (Link: <u>Sample corporate policy for use</u> <u>of Generative Artificial Intelligence - Socitm</u>) and used by other local authorities, such as Hounslow and Bristol City Councils, as a basis for their AI Policies.
- 34 The various governance groups have been working to produce the GenAI Policy:
 - draft briefing paper to SIGG for comment and approval
 - briefing paper to Senior Responsible Officer, Digital Workstream, for information and comment.
 - SIGG AI Task and Finish Group established enabling consultation with stakeholders including adults and children's services, public health, finance, legal and human resources.
 - draft GenAl policy presented to SIGG, Corporate Services Leadership Team (CSLT), Digital Workstream and Corporate Leadership Team (CLT) for comment and endorsement.

Implications

Monitoring Officer/Legal

- 35 Legal worked with colleagues on the SIGG Task and Finish Group to draft the GenAI Policy and duly considered (and not limited to) the following:
 - (a) Consideration must be given to the fact that, without appropriate controls, data entered into GenAI may enter the public domain. This can release non-public information and breach regulatory requirements, customer, or vendor contracts, or compromise intellectual property. Users will need to be made aware of the GenAI Policy to ensure that their use of any GenAI complies with all applicable laws and regulations and with associated Council policies.

CoPilot is adopted as the standard tool for GenAI as this provides an area where user and business data is protected and will not leak outside the Council. Chat data is not saved, and it is not used to train other AI systems.

(b) It clearly states that employees must not use or enter confidential, proprietary, or trade secret information into unauthorised AI chatbots or language models, such as ChatGPT.

- (c) Staff will also need to follow all applicable data privacy laws and Council policies when using GenAI. Any release of private/personal information without the authorisation of the information's owner will result in a breach of relevant data protection laws. If a user has any doubt about the confidentiality of information, they should not use GenAI.
- (d) Consideration has been given to copyright laws when using GenAI. It is prohibited to use GenAI to generate content that infringes upon the intellectual property rights of others, including but not limited to copyrighted material. It is likely that if a user is unsure whether a particular use of GenAI constitutes copyright infringement, they will be advised to contact their legal advisor or Information Governance Collaborative Group before using GenAI.
- (e) All information generated by GenAI must be reviewed and edited for accuracy prior to use. Users of GenAI are responsible for reviewing output and are accountable for ensuring the accuracy of GenAI generated output before use/release. If a user has any doubt about the accuracy of information generated by GenAI, they should not use GenAI.
- 36 Underpinning security and compliance have been considered. While an GenAl platform may be hosted internationally, information created or collected in the United Kingdom of Great Britain and Northern Ireland (UK), under data sovereignty rules, is still under jurisdiction of UK laws. The reverse also applies. If information is sourced from GenAl hosted overseas for use in the UK, the laws of the source country regarding its use and access may apply. GenAl service providers should be assessed for data sovereignty practice by any organisation wishing to use their GenAl.

Section 151 Officer/Finance

37 There are no financial implications of this briefing report and its recommendations/decisions. Any financial implications associated with GenAI developments will be documented and presented as business cases through the Digital Workstream. If these business cases are approved, they will be drawn down from the Council's approved Digital budget as part of the Medium-Term Financial Strategy (MTFS).

The Digital Workstream is also considering what change would look like (for example, % of services which will include AI processes, number or percentage of interactions that will be AI supported, the availability of extended hours, the alignment of data, the reduction in licenses and staffing requirements). Preliminary contact statistics (01 Apr 2022 – 31 March 2023) being considered for AI solutions and automation include:

- Face to face reception requests approx. 13405 (Macclesfield (5424) and Crewe (7981)).
- Phone calls of which 326,574 are to the contact centre alone (non-Contact Centre Calls will also to be investigated).
- Contact Centre emails of approx. 600,415 (which include outbound and inbound requests).
- Council emails receives approx. 22,959,976 and sends approx. 19,561,550.
- Online forms completed approx. 191,096.
- Postal requests are being investigated (estimate + 1m), and
- Cheshire East Web Site Contacts approx. 10,000,000.

Initial AI projects therefore being scoped by the Digital Workstream include the following and will identify the number or percentage of interactions that will be AI supported, the availability of extended hours, the alignment of data, the reduction in licenses and staffing requirements:

- AI email and voice automation for the Contact Centre, Highways, Waste and linked services Communities and Street Cleansing services.
- Generative Chat for all users
- Web Journeys linked to a Council AI knowledge base for the Contact Centre, Highways, Waste and linked services Communities and Street Cleanse initially.

Policy

- 38 Al is expected to deliver significant benefit for the UK population, it is booming in healthcare and biology and may help the Council mitigate the impact of climate change.
- 39 Al systems have been commercially used to automate or streamline internal processes, such as classifying customer support requests to route them to the correct department, filling out or parsing forms, detecting anomalous behaviour like fraud or cyber intrusions, or screening, hiring, or evaluating employees.
- 40 The Council may therefore want to use AI within a service to provide services both internally and externally, or an officer may use it to provide research and guidance. Benefits may include:
 - The ability to summarise or extract meaning from existing documents, policies, or systems.
 - The provision of a service externally to residents instead of human intervention

- Providing access to resources in the public domain for advice and guidance internally, or
- Exposing information externally to enable transparency.
- 41 There are global, regional, and local authority examples for GenAl being deployed in support of, for example:
 - (a) City planning, economic regeneration, and data infrastructure agency opportunities
 - (b) Smart Cities
 - (c) Geospatial data intelligence, and 3D mapping data
 - (d) Cross-agency cyber command and cyber defence applications
 - (e) Administrative Services, energy billing, and robotic process automation for example, email and telephone response assistants, chatbots, and appointments
 - (f) Environmental protection, health and mental hygiene, and noise data analytics
 - (g) High school admissions, matching, and other algorithmic approaches
 - (h) Education and learning
 - (i) Community engagement and participatory approaches
 - (j) Foster partnership opportunities with external experts
 - (k) Population health studies
 - (I) Clinical interventions, and
 - (m) Food poverty and participatory system design all or some of which the Council may want to consider.
- 42 The Council is scheduled to implement the following Digital Customer Enablement GenAl projects in 2023-4 for Highways specifically (a to c) and corporately (d to f):
 - (a) Sign Posting and Streamlining of Reporting Channels such as:
 - Enquiries for other Council services or other third parties (Town and Parish Councils / Housing Associations etc). Last year, 7% of our enquiries were not for the Highway Service – equating to 2,200 enquiries.
 - (ii) Looking at how can get these to the right service / external organisation at the right time?
 - (b) Recognising an 'emergency' when a customer is logging an enquiry a report:
 - (i) Anything considered as an emergency we would want to direct the customer through to reporting the enquiry via telephone via the emergency line 0300 123 502.
 - (ii) If the customer uses certain words (fallen over / injured myself / tyre blown etc.) the system will recognise this and send the customer via the telephone route?
 - (c) Information Requests providing intelligent responses at point of contact, for instance:

- (i) Frequently asked questions that could be resolved at first point of contact by an AI solution looking at current programmes on the Council's website, for instance:
 - 1. When are you going to cut the grass? Identifying who is responsible for the ownership / maintenance of the grass reported also.
 - 2. When are you going to repair my road?
- (d) Automation of request and delivery processes for example:
 - (i) Logging a pothole on Fix My Street:
 - 1. Customers can currently log anything that they consider as a pothole with limited guidance within the tool. The guidance needs to be reviewed with respect to what can / or could be logged.
 - 2. Would be good to advise the customer whether there has been a recent safety inspection and whether any defects identified?
 - 3. Do we need customers to log duplicate enquiries if we are already aware?
- (e) Email Automation Avoidable Processing
 - (i) Filtering of Duplicate Mails
 - (ii) Automated Service Request
 - (iii) Automated Information Request
 - (iv) Re-Direction of Mails
- (f) Telephony & Voice Automation Avoidable Contact
 - (i) Sign Posting
 - (ii) Automated Service Request
 - (iii) Automated Information Request
 - (iv) Re-Direction of Calls and Call Backs
- 43 The Council has access to Microsoft's Bing Chat Enterprise and CoPilot.
- 44 CoPilot integrates a LLM with the Microsoft set of tools. It differs from a public LLM in that the LLM respects organisational barriers. Information posted and replied to comes from within that barrier, enabling safe use within the constraints on normal information protection.
- 45 Within the press, experts are therefore raising concerns that AI is not necessarily too powerful, but that the basic guard rails are not in place to ensure AI is deployed responsibly.
- 46 The Information Commissioners Office (ICO) has written specific guidance and additions that must be in place. AI must be considered as part of the Data Protection Impact Assessment (DPIA) process which helps to identify and minimise the data protection risks of the AI being implemented. This needs to describe the nature, scope, context, and

purposes of any processing of personal data. It needs to make clear how and why the Council are going to use AI to process the data. We will need to detail:

- (a) how we will collect, store, and use information.
- (b) the volume, variety, and sensitivity of the information.
- (c) the nature of your relationship with individuals; and
- (d) the intended outcomes for individuals or wider society, as well as for the Council.
- 47 Whether a system using AI is generally more or less risky than a system not using AI depends on the specific circumstances. The DPIA should show evidence of our consideration of less risky alternatives, if any, that achieve the same purpose of the processing, and why we did not choose them. This consideration is particularly relevant where we are using public task or legitimate interests as a lawful basis. The DPIA will need to be reviewed to incorporate AI and its use. In addition, to retain the trust of the public, business users will need to consider the use of AI from a privacy and transparency basis with clear notice on where information and decision have used these AI tools.
- 48 Fundamentally, any policy is only as good as the governance framework that underpins it. In the context of AI, governance around data use and sharing is of particular importance, so additional thought must be given to governance models that facilitate access to large data sets to improve the quality of data input and the integrity of outcomes, while ensuring patients or cared-for customers have control over how their data is used, who has access to it and for what purposes.
- 49 While values and principles are crucial to establishing a basis for any ethical AI governance framework, recent movements in AI ethics have emphasised the need to move beyond high-level principles and toward practical strategies.
- 50 The Strategic Information Governance Group (SIGG) advised the Digital Workstream on the governance framework and policies that underpin the development of AI solutions across the Council. The draft GenAI Policy in Appendix 1 recommends governance mechanisms which are inclusive, transparent, multidisciplinary, multilateral, and multistakeholder. In other words, communities impacted by AI must be actively involved in its governance in addition to experts across a range of disciplines. Additionally, governance must extend beyond mere recommendations to include anticipation, enforcement, and redress. AI governance cannot stop once risks and impacts have been identified. Instead, all identified harms must be investigated and addressed so that impacted communities have the right to redress.

- 51 As SIGG has the responsibility for corporate data governance, it will implement mechanisms for effective AI data governance strategies to ensure individual privacy while ensuring adequate data collection and means to regulate its use.
- 52 GenAI may store sensitive data and information, which could be at risk of being breached or hacked. The Council must assess technical protections and security certification of GenAI before use. If a user has any doubt about the security of information input into GenAI, they should not use GenAI. Any procurement of GenAI software must adhere to current procurement processes including the need for a completed and approved security questionnaire.
- 53 All policies, processes and procedures will need to be reviewed periodically and updated as necessary to ensure continued compliance with all applicable legislation, regulations, and wider Council policies.
- 54 The draft GenAl Policy supports the effective and enabling organisation referred to within the priorities in the <u>Corporate Plan</u> that this briefing report supports.

Equality, Diversity, and Inclusion

- 55 Ethics is about more than laws and regulations, compliance, and checklists. It is about designing the world we want to live in. The use of AI raises a series of ethical issues that academics and policymakers have grappled with for years, and which local government embracing the technology must address also.
- 56 Al technology brings major benefits in many areas, but without ethical guardrails, it risks reproducing real world biases and discrimination.
- 57 Use of GenAI must be in a manner that promotes fairness and avoids bias to prevent discrimination and promote equal treatment and be in such a way as to contribute positively to the Council's goals and values.
- 58 An ethical concern surrounding AI is the risk of it being used to spread misinformation. As it is designed to generate responses that sound human-like, it can be difficult for people to distinguish between responses generated by a machine and those that come from a real person. It can also be seen that regardless of whether a response is known to emanate from an LLM people are affected by that response.
- 59 GenAl must be used ethically and in compliance with all applicable legislation, regulations, and Council policies. Users may advise that they should not use GenAl to generate content that is discriminatory, offensive, or inappropriate. If there are any doubts about the appropriateness of using GenAl in a particular situation, users will be

advised to consult with their line manager or Information Governance Collaborative Group.

- 60 The European Commission published "The final assessment list for trustworthy AI" in July 2020 and translates the ethics guidelines into an accessible checklist that developers and deployers of AI can use, touching on seven requirements deemed essential to the ethical use of AI:
 - human agency and oversight.
 - technical robustness and safety.
 - privacy and data governance.
 - transparency.
 - diversity, non-discrimination, and fairness.
 - societal and environmental wellbeing; and
 - accountability.
- 61 Without these principles it can lead to the dissemination of false or misleading information, which can have profound consequences. For example, if a chatbot is trained on biased or outdated data, it may produce responses that reflect that bias, potentially leading to the spread of misinformation.
- 62 Areas that are already safety-critical or covered by strict product legislation such as medical devices, are already required to undergo conformity assessment. This could be a reason the use of AI in healthcare is so advanced. Social Care works with healthcare providers and may need to understand their legal and regulatory obligations and have an appreciation of the ethical guidance that exists, and associated policy that allows the organisation to meet the requirements and expectations in practice. For instance, AstraZeneca has a set of its own principles for ethical data and AI to guide the approach of its staff.
- 63 There is significant cross-over between the AstraZeneca principles and the ethics guide and checklist endorsed by the European Commission. The principles revolve around five core themes:
 - Explainable and transparent.
 - Fair.
 - Accountable.
 - Human-centric and socially beneficial; and
 - Private and secure.
- 64 Among the specific commitments the company outlines, AstraZeneca promises to be "open about the use, strengths and limitations of our data and AI systems," to ensure humans oversee AI systems, to ensure data and AI systems are secure, and to "act in a manner compatible with intended data use." It also states that it anticipates and mitigates

the impact of potential unfavourable consequences of AI through testing, governance, and procedures, and further promises to learn lessons from "unintended consequences" materialising from its use of AI.

- 65 Novartis has also developed its own <u>ethical principles for Al</u> <u>development, application and use</u> too. In a detailed document, the company has, among other things, specifically acknowledged the risk of bias and discriminatory outcomes from using unrepresentative data samples.
- 66 <u>Sanofi is another company in the sector developing its own policy on the</u> <u>use and governance of AI</u>. It has said it will be shaped around three principles:
 - Al should be used in the interest of patients.
 - The use of AI should not treat any groups of patients unfairly.
 - Dignity needs to be preserved so the patient should have autonomy of thought, intention, and action when making decisions regarding health care.
- 67 The Department for Education has produced guidance on the use of generative artificial intelligence (AI), including large language models (LLMs) like ChatGPT or Google Bard, in the education sector <u>Generative artificial intelligence in education - GOV.UK (www.gov.uk)</u>. It is informed by Government's White Paper on a pro-innovation approach to AI regulation and follows Government's recent announcement to create a Foundation Model Taskforce which will consider UK domestic capability in this important technology. Generative AI refers to technology that can be used to create new content based on large volumes of data that models have been trained on. This can include audio, code, images, text, simulations, and videos.
- 68 Artificial Intelligence (AI) has the potential to address some of the biggest challenges in education today, innovate teaching and learning practices, and accelerate progress towards SDG 4. UNESCO, for instance, is committed to supporting Member States to harness the potential of AI technologies for achieving the Education 2030 Agenda, while ensuring that its application in educational contexts is guided by the core principles of inclusion and equity. Furthermore, UNESCO has developed within the framework of the Beijing Consensus a publication aimed at fostering the readiness of education policymakers in artificial intelligence. This publication, Artificial Intelligence and Education: Guidance for Policymakers, will be of interest to practitioners and professionals in the policymaking and education communities. It aims to generate a shared understanding of the opportunities and challenges that AI offers for education, as well as its implications for the core competencies needed in the AI era.

- 69 Addressing ethical issues must be at the heart of building trust in the use of AI in social and health care. The body driving digital transformation in healthcare in England, NHSX, is taking its own steps to address the ethical challenges posed by AI use in healthcare. It has set up an NHS AI Lab and its AI ethics initiative seeks to "ensure that AI products used in the NHS and care settings will not exacerbate health inequalities."
- 70 SIGG will also look at an Ethical Impact Assessment (EIA) which provides an opportunity to reflect on the potential impacts of an AI project and to identify the needed harm prevention actions. The EIA will address questions such as: Who is most likely to be adversely affected by this AI system? What form will these impacts take? What can be done to prevent these harms and have resources been allocated to this harm prevention?

Human Resources

- 71 Human Resources are represented at SIGG and were consulted during the development of the draft GenAl Policy in Appendix 1.
- 72 Users may want to use GenAl for work-related purposes subject to adherence to the policy. This includes tasks such as generating text or content for reports, emails, presentations, images, and customer service communications.

Risk Management

- 73 Issues range from how the use of AI may fundamentally alter the nature of work undertaken by humans, to how to address the risk that advances in machine learning creates systems that become too smart for humans to control. There are further ethical questions around what happens, and who is responsible, when AI systems make a mistake; how the use of AI can respect privacy and freedom of expression; and how to ensure there is not inherent bias in the way AI systems function that might deliver inaccurate, discriminatory, or even dangerous outcomes.
- 74 The problems with such systems are not always evident in common use, but in one example, ChatGPT has been found to advise patients with mental health issues to die by suicide. Many other examples of unsafe responses can be found and envisaged, and it will be impossible to prevent these without significant development in the state of the art. This is even acknowledged by ChatGPT itself, which says "... ChatGPT, like other language processing models, can generate biased or offensive text if it is trained on biased data. This can be a problem if the

generated text is used in sensitive applications, such as in healthcare or education, where it is important to avoid bias and promote inclusivity."

- 75 Chatbots are therefore in their infancy and need sufficient prompts within questions to provide meaningful responses. Without the right context then often, the responses are misleading or even wrong, so care needs to be taken deriving decisions from content. When used to support customers, this may lead to frustrations and inaccurate signposting.
- 76 Any use of GenAl technology in pursuit of Council activities should be done with full acknowledgement of the policies, practices, terms, and conditions of developers/vendors.
- 77 Information Governance policies such as Data Protection state that staff must not disclose sensitive or confidential data (e.g., into public LLM's as this information will then be freely available to other chatbots). The Al policy will need to reflect other council policies such as Information Governance and HR policies.
- 78 Employees may also receive information through these chatbots that is trademarked, copyrighted, or the intellectual property of another person or entity, creating legal risk for the council as the employer.

Rural Communities

79 There are no implications at this stage to rural communities of this report and its recommendations.

Children and Young People including Cared for Children, care leavers and Children with special educational needs and disabilities (SEND)

- 80 Adults and Childrens Services SIGG representatives have been consulted in the development of the draft GenAI Policy Appendix 1.
- 81 It should be noted that this briefing report, and the draft GenAI Policy, does not include Schools.
- 82 UNESCO's mandate calls inherently for a <u>human-centred approach to</u> <u>AI</u>. It aims to shift the conversation to include AI's role in addressing current inequalities regarding access to knowledge, research, and the diversity of cultural expressions and to ensure AI does not widen the technological divides within and between countries. The promise of "AI for all" must be that everyone can take advantage of the technological revolution under way and access its fruits, notably in terms of innovation and knowledge.

- 83 Although generative AI is not new, recent advances and public access to the technology mean that the public can now use this technology to produce AI-generated content. This poses opportunities and challenges for the education sector, children and young people, care leavers and Children with special educational needs and disabilities (SEND):
 - (a) When used appropriately, technology (including generative AI), has the potential to reduce workload across the sector, and free up time, allowing users to focus on delivering excellent services.
 - (b) Those responsible for children and young people, care leavers and children with special educational needs will need to continue to take reasonable steps (where applicable) to prevent malpractice, including malpractice involving use of generative AI and other emerging technologies.
 - (c) The education sector must continue to protect its data, resources, staff, and pupils, in particular:
 - (i) Personal and sensitive data must be protected and therefore must not be entered into generative AI tools.
 - (ii) Education institutions should review and strengthen their cyber security, particularly as generative AI could increase the sophistication and credibility of attacks.
 - (iii) Education institutions must continue to protect their students from harmful content online, including that which might be produced by generative AI.

Public Health

- 84 Adult Social Care and Public Health SIGG representatives have been consulted in the development of the draft GenAI Policy, including the following areas:
 - Digital inclusion issues and opportunities
 - Any positive, neutral, or negative overall impacts on the health and wellbeing of Cheshire East residents, and
 - Any greater (positive or negative) impact on some groups compared to others (e.g., rural vs urban; younger vs older; poorer vs more affluent; etc.)

Climate Change

85 Approved GenAI projects will document any potential opportunities to reduce the Council's carbon footprint and achieve environmental sustainability.

Access to Information	
Contact Officer:	Contact Officer: Gareth.Pawlett@cheshireeast.gov.uk
Appendices:	Appendix 1: Draft GenAl Policy
Background Papers:	NA