

THEME A: CONCEPTS · A4.4

Ethical considerations — Mark scheme

33 marks · covers A4.4.1–A4.4.2

How to mark. Each creditworthy point is on its own line and ends with a semicolon (;). A line beginning with ... is an *expansion* of the line above it, so the pair is worth the marks shown. A slash (/) separates acceptable alternative wordings. Words in (brackets) are not needed for the mark. Mark positively: credit what the student has got right rather than penalising what is missing, and accept any valid alternative not listed here. Do not award more than the maximum for a part. Where a later part uses a wrong answer from an earlier part correctly, award follow-through marks (FT). HL marks parts that assess higher level only content.

Q	ANSWER	MARKS
1(a) A4.4.1	Award [3] max Award [1] per point up to [3] max. Must be linked to this training data. The model learns the patterns present in the past decisions it was trained on; Because few women were hired into technical roles, the data encodes that outcome as the normal result; ...(so) the model is likely to score female applicants lower and shortlist fewer of them, whatever their qualifications; The model may also pick up on proxies for gender, such as the name of a school or a sport, even if gender itself is removed; The historical bias is therefore repeated and given the appearance of an objective, data-driven decision;	3
1(b) A4.4.1	Award [2] max Award [1] for the point and [1] for the expansion. Any one method. Rebalance the training data so that successful applicants of each group are represented fairly; ...(so) the model is not learning from a skewed record of past decisions; Audit the model's outputs by group and measure the shortlisting rate for each; ...(so) an unfair disparity is detected before the model is used on real applicants; Keep a human in the loop to review the model's decisions; ...(so) a rejection can be challenged and overturned rather than being final; Remove attributes that act as proxies for gender; ...(though) this alone is not sufficient, since other attributes can still correlate with it;	2
1(c) A4.4.1	Award [1] max The principle that a model's decisions should not systematically advantage or disadvantage people because of a protected characteristic such as gender, race or age;	1
2(a) A4.4.1	Award [3] max Award [1] per point up to [3] max. Transparency allows applicants to know how a decision about them was reached; An applicant cannot challenge or appeal a decision whose basis they do not know about; They may have a legal right to be informed about automated decision-making, and to request human review;	3

Q	ANSWER	MARKS
	Knowing the process is automated allows applicants to judge whether it is appropriate and to raise concerns about it; Concealing it damages trust in the company if it later becomes known;	
2(b) A4.4.1	Award [3] max This is a discuss question (AO3), so more than one position must be considered. Award [1] per point up to [3] max. Award a maximum of [2] if only one party is considered. The company that deployed it chose to use the model and benefits from it, so it carries responsibility for the outcomes; It also chose the training data and set the threshold at which applicants are rejected; The developers built the model and should have tested it for bias before release, so they share responsibility; ...(though) they may not control how the company later applies it; Arguments about diffused responsibility No single person made the decision, which is precisely the problem: automation can leave accountability unclear; Blaming the model itself is not meaningful, since software cannot be held responsible; Conclusion (credit any supported judgement, for example) Accountability should rest with the company, since it decided to deploy the system and to act on its output, with a duty on developers to disclose the model's known limitations;	3
3(a) A4.4.1	Award [2] max Award [1] per point up to [2] max. The user must be told clearly what data is collected, who it is shared with and what it will be used for, before agreeing; The explanation must be in plain language rather than buried in a long terms-of-service document; Consent must be freely given and refusible, so the app should still work if the user declines the sharing; The user must be able to withdraw consent later;	2
3(b) A4.4.1	Award [3] max Award [1] per point up to [3] max. Any one risk, developed. Health and location data is highly sensitive and can reveal medical conditions the user has not disclosed; The insurer could raise premiums, or refuse cover, based on inferences the user never agreed to be judged on; ...(so) users may be penalised for patterns in their data rather than for anything they have actually claimed for; Users may change their behaviour because they know they are being monitored; Once the data is transferred, the user has little control over how long it is kept or who else it reaches; A breach at either company would expose intimate details about the user's health and movements;	3
4(a) A4.4.1	Award [2] max Award [1] per point up to [2] max. Training runs for long periods on large numbers of processors, consuming a great deal of electricity;	2

Q	ANSWER	MARKS
	...(so) it produces significant carbon emissions unless the data centre runs on renewable energy; Data centres also consume large volumes of water for cooling; Manufacturing and replacing the specialised hardware has its own environmental cost;	
4(b) A4.4.2	Award [4] max This is a discuss question (AO3), so both benefits and harms are required. Award [1] per point up to [4] max, but award a maximum of [3] if only one side is given. Arguments that it benefits society Work that is dangerous, repetitive or physically damaging can be taken away from people; Goods and services become cheaper and more widely available; Productivity rises, which can fund shorter working hours or better public services; New categories of work are created, as has happened after previous waves of automation; Arguments that it harms society People lose their income, and the loss falls on specific communities rather than being spread evenly; The new jobs created often require different skills and appear in different places, so displaced workers cannot simply move into them; The gains flow mainly to whoever owns the technology, widening inequality; Work carries meaning and identity beyond income, so the harm is not purely financial; Conclusion (credit any supported judgement, for example) Whether automation benefits society depends less on the technology than on whether retraining and income support are put in place alongside it;	4
5(a) A4.4.1	Award [3] max Award [1] per point up to [3] max. The model optimises for engagement, not for whether a post is true; Content that provokes outrage, fear or surprise holds attention, and false claims are often more striking than accurate ones; ...(so) the model learns to promote such posts because they keep users on the platform longer; Users are shown more of what they already engage with, creating a feedback loop that narrows what they see; The scale of automated recommendation means a false claim can reach millions before any human reviews it;	3
5(b) A4.4.1	Award [2] max Award [1] for the point and [1] for the expansion. Any one measure. Include a signal for source reliability in the ranking, not only engagement; ...(so) content from unreliable sources is shown less even when it is engaging; Have disputed posts checked by human moderators or fact-checkers; ...(so) claims flagged as false can be labelled or demoted; Limit how quickly a post can be reshared; ...(so) there is time for a false claim to be reviewed before it spreads widely;	2
6 A4.4.2	Award [5] max This is a discuss question (AO3). The question is about the privacy of bystanders, not of the wearer. Award [1] per point up to [5] max, award a maximum of [3] if only one side is given, and a maximum of [4] without a supported conclusion.	5

Q	ANSWER	MARKS
	Concerns for people around the wearer Bystanders are recorded without being asked and often without realising it, so they cannot give or refuse consent; Facial identification means a stranger can learn who someone is, which removes the practical anonymity of public space; Recordings could reveal where a person was, when, and who they were with, building a movement history over time; People behave differently when they suspect they are being recorded, which has a chilling effect on ordinary conversation; The risk falls unequally: it is greater for people who are already vulnerable, such as those escaping an abusive partner; The recordings are stored by a company that the bystander has no relationship with and cannot ask to delete them; Arguments on the other side There is an established expectation of reduced privacy in genuinely public places; The glasses offer real benefits, for example to people with visual impairments or memory difficulties; Recording can protect people by providing evidence of crimes or of abuse of authority; Regulation, indicator lights, on-device processing and automatic face blurring could reduce the harm without banning the technology; Conclusion (credit any supported judgement, for example) The central problem is that the cost falls on people who receive none of the benefit and never agreed to it, so the technology needs enforceable limits rather than reliance on users behaving well; This is why ethical guidelines must be reassessed as new technologies appear, rather than being settled once;	

Total for this paper: 33 marks