# Report and Recommendations
# 'Alice'
# Ethical Considerations for AI in Education

Enzo Joly, 22055453

Module Title: UFCFB5-15-3 | Ethical and Professional Issues in Computing and Digital Media
Word Count: 2461

What ethical or social issues are important for this project?
What professional, legal or design standards might the organisation need to read and implement?
What can YOU suggest as KEY actions to take to ensure that this project is successful?

# Contents

# 1 Introduction

This report comprises ethical considerations and analysis concerning technology, data privacy, disinformation, and accessibility.

The resulting recommendations for deployment aim to inform, such that TechSoft be fully cognisant of challenges faced by an AI chatbot in education. South Star Academy's proposed student support services chatbot 'Alice' will require thoughtful diligence before it can be deemed production-ready.

# 2 Relevant Codes of Practice

Professional codes of conduct provide software engineers with an ethical starting point for a deployment of this nature. Given the bandwidth of this topic, and the scope of this report, exhaustive coverage of the full documentation is impossible, so key characteristics of differing guidelines will be the primary focal point. The BCS focusses on professionals; the IEEE standardises the technological aspects; the ACM is more concerned with societal impact.

## 2.1 BCS Code of Conduct

BCS members must demonstrate "due regard for public health, privacy, security and wellbeing of others and the environment" *(British Computer Society 2024, p. 2)*, particularly salient considering Alice will be handling sensitive student data and potentially influencing student wellbeing.

The BCS also mandates professionals to only provide service "within [their] professional competence" *(British Computer Society 2024, p. 2)* which begs the question: where exists the limit of Alice's 'competence' in the context of student support? Is a system's competency assumed to be an extension of the professionals who calibrated it?

Members are additionally required to "uphold the reputation of the profession" *(British Computer Society 2024, p. 3)*. This translates to full transparency regarding Alice's management and capabilities. Safeguards must also be in place to prevent misuse in order to preserve public trust of AI systems in education, let alone the trust of students.

## 2.2 IEEE Code of Ethics

Data retention and deletion policies must be established and clear, ensuring a commitment to "protection of... personal information and data" *(Institute of Electrical and Electronics Engineers 2024, p. 1)*, and a commitment to student privacy. Engineers must identify and eliminate potential bias from the model and algorithms in use, such that Alice "not discriminate against any person because of characteristics protected by law" *(Institute of Electrical and Electronics Engineers 2024, p. 1)*. This should be verified and assured with rigorous testing and regular audits, ensuring commitment to "honest and realistic... claims or estimates based on available data" *(Institute of Electrical and Electronics Engineers 2024, p. 2)*, coupled with an established mechanism for human oversight and intervention when Alice's confidence in a response is low.

## 2.3 ACM Code of Ethics

The ACM Code proposes that "all people are stakeholders in computing" *(Association for Computing Machinery 2024, p. 1)*, alluding that the enhancement of student wellbeing and academic success is not merely a nice gesture, but necessary for the betterment of society, assuming compliant (and considerate) execution.

Alice should be constructed with accessibility in mind, "[fostering] fair participation of all people, including those of underrepresented groups" *(Association for Computing Machinery 2024, p. 2)*.

#### Caveats to Information Processing

Blundell argues that "such codes are seldom consulted and often incorporate bland (and sometimes contradictory) statements intended to satisfy a broad range of stakeholders" *(Blundell 2020, p. 40)* and professionals must practise discernment and maintain awareness for the consequences of poor decision-making, both quantitatively and in qualitative aspects. Underpinning deontological ethics is the notion that the professional is duty-bound to rely on their better judgement where universally established rules fail. This remains one of Kant's core philosophical tenets, to "respect the reason in you".

BCS specifies duty to "due... diligence in accordance with the Relevant Authority's requirements whilst exercising... professional judgement at all times" *(British Computer Society 2024, p. 2)*. In our context: information relating to the endangerment of any student(s) would likely take precedent over privacy, as is established protocol *In Loco Parentis*. This should be at the discretion of human moderators by way of a sophisticated content-flagging system and anonymity secured by multiple key-holders to ensure consensus for de-anonymisation. Students may be more comfortable anonymously, so this process of de-anonymising must be abuse-proof to ensure trust remains strong. All decision-making processes must prioritise students foremost, with the necessary privacy precautions in place to protect student data alongside regular monitoring of the chatbot's holistic impact on the educational environment.

# 3 Further Considerations

**A Note on Regulatory Standards**

In-depth representation for truly ethical and compliant practice are appended in Appendix B (relevant definitions with respect to implications for Alice), and not the main body of this report, for brevity's sake.

## 3.1 Ethical Data Governance

Any implementation of 'Alice' would raise significant privacy concerns regarding the collection, storage, and use of student data *(Annus 2023, pp. 366-370)*. Since proper compliance is in the best interest of student wellbeing and vice versa, all-encompassing singular actions, policies, and processes have the potential to satisfy all stakeholders simulataneously.

### 3.1.1 GDPR/Regulatory Compliance

From design through to deployment continously for any collection of data, as well as full transparency in its use of data, and clarity regarding what it is collecting data to be used for, TechSoft must define the nature of its policies, processes, and practice. It is also imperative that TechSoft safeguard against any third-parties from gaining access to data with security best practices. Schools are in position to safeguard student data, appreciating this perspective is essential for TechSoft engineers. For more information, see Appendix B. Additionally, record a Data Protection Impact Assessment (DPIA) (template provided in Appendix A).

### 3.1.2 Data Security

Privacy by design as an ethos enabling strong governance and security measures including:

**Authentication and Authorisation**

OAuth 2.0 and OpenID Connect for Single Sign-On (SSO) and Role-Based Access Control (RBAC) for fine-tuning permissions *(Josuttis 2023, pp. 80-120)*.

**Data Encryption**

Robust data encryption methods *(Stallings 2023, pp. 100-150)*, AES-256 for data at rest and TLS 1.3 for data in transit.

### 3.1.3 Data Anonymisation

K-anonymity can be employed for protecting student identities *(El Emam and Arbuckle 2023, pp. 75-100)*, and the implementation of differential privacy for aggregate data analysis, anonymising students entirely into statistics beneficial for upgrading and improving the chatbot, and alleviating the burden of at-risk data unnecessary for development purposes.

### 3.1.4 Data Minimisation

Data "audited regularly to ensure all stored information is still relevant" *(Article 29 Working Party 2018)*, implementing "privacy-preserving UX patterns" *(Hartzog 2023, pp. 50-100)*, making privacy settings easily accessible and understandable.

**Informed Consent**

Information about data usage must be clear, with opt-in mechanisms for non-essential features. Age-appropriate consent must be obtained from students (parental consent for students under the age of 13 *(Federal Trade Commission 2023)*).

## 3.2 AI Ethics Principles

Adhering to established AI ethics principles *(European Commission 2024)*.

### 3.2.1 Ethical Algorithm Design

Explicit labelling of 'Alice' as an AI system *(Institute of Electrical and Electronics Engineers 2023)*, including information on the capabilities and management of the chatbot itself. Elucidation of empirical evidence rooting from the consequences of algorithmic configuration choices can benefit not only developers, but users, and by extension society en masse.

### 3.2.2 Human Agency and Oversight

Setting "clear boundaries between AI support and human intervention" *(American Psychological Association 2024)*by defining thresholds for transitioning between human support and AI. Staff training on effectively working alongside AI systems for smooth integration of technology. This will also enable contextual enquiry to better map student support scenarios, conducting user research *(Goodman, Kuniavsky, and Moed 2023, pp. 50-100)*, and human-in-the-loop dialogue optimisation *(Vaughan 2024, pp. 30-60)*. Evaluations can be crowdsourced for diverse perspectives.

#### Iterative Design Process

Implementing an iterative design process *(Holtzblatt and Beyer 2024, pp. 30-60)* will aid in refactoring and improving an adaptable system to student needs, and usability testing with representative student groups will ensure that these needs are measured and met.

#### Ethical Review Process

Implementing an ethical review process *(Floridi and Cowls 2023)* includes periodic assessments of chatbot decisions for bias, and alignment checks with established ethical guidelines.

### 3.2.3 Technical Safety

Implementing safeguards against harmful or inappropriate chatbot responses *(Bickmore et al. 2021, p. e11510)*, safeguarding against jailbreaking and exploitation by students to ensure integrity of all responses. Content filtering and trigger warning systems as well as established escalation protocols for crisis situations. Natural and effective chatbot interactions, establishing tone and personality *(Bradbury 2024, pp. 20-50)*, maintaining a consistent voice aligned with an educational context, ensuring age-appropriate content.

### 3.2.4 Transparency

"Explainable AI techniques to interpret chatbot decisions" *(Arrieta et al. 2022, pp. 82-115)* by way of interpretable machine learning models and provision of rationale, with respect to design recommendations in the software lifecycle.

### 3.2.5 Fairness-Aware Machine Learning

Implementing fairness-aware machine learning techniques *(Barocas, Hardt, and Narayanan 2021)* is crucial:

- Utilising algorithmic fairness metrics (e.g., demographic parity, equal opportunity)
- Applying bias mitigation strategies in model training and deployment

#### Bias Testing

Mitigating bias in the chatbot's algorithms is essential for equitable student support: Regular testing for biases in chatbot responses *(Association for Computing Machinery Conference on Fairness, Accountability, and Transparency 2024)* should include:

- Employing automated bias detection tools
- Implementing human-in-the-loop evaluation for sensitive topics

#### Diverse Training Data

Using diverse training data to prevent demographic biases *(Mehrabi et al. 2023, pp. 1-35)* involves:

- Including diverse student profiles in training datasets
- Regularly updating data to reflect changing student demographics

### 3.2.6 Regular Audits

Conducting regular audits of the chatbot's decision-making processes *(AI Ethics Guidelines for Education 2024)* is essential:

- Logging and analysis of chatbot interactions
- Third-party audits to ensure adherence to ethical guidelines

## 3.3 User Experience and Accessibility

### 3.3.1 User-Centred Design

Implementing a design process focussed on student needs:

### 3.3.2 Accessibility Standards

Ensuring the chatbot is inclusive of all students:

Adhering to WCAG 2.2 *(World Wide Web Consortium 2023)* includes:

- Ensuring content is perceivable, operable, understandable, and robust
- Implementing keyboard accessibility and enough time for user interactions

#### Cognitive Accessibility Considerations

Addressing cognitive accessibility *(Yesilada, Brajnik, and Harper 2023, pp. 1-10)* includes:

- Using clear and simple language, multi-lingual options *(Anastasiou and Schäler 2023, pp. 50-100)*.
- Implementing consistent layout and interaction patterns

#### Adaptable User Interface

Creating an adaptable user interface *(Harper and Yesilada 2024, pp. 20-50)* includes:

- Offering customisable font sizes and colour contrasts
- Supporting different input methods (text, voice, gestures)

#### Neurodiversity Considerations

Addressing neurodiversity *(Armstrong 2023, pp. 30-60)* involves:

- Providing options to reduce visual clutter
- Offering alternative formats for information presentation (text, audio, visual)

#### Touch-Friendly Interfaces

Designing touch-friendly interfaces *(Hoober and Berkman 2023, pp. 80-120)* involves:

- Using appropriately sized touch targets
- Implementing gesture-based interactions where appropriate

### 3.3.3 User Feedback Integration

Incorporating student and staff input:

#### Feedback Collection Methods

Implementing feedback collection methods *(Tullis and Albert 2024, pp. 100-50)* involves:

- Providing in-chat feedback options
- Conducting periodic user surveys

#### In-App Feedback Mechanisms

Implementing in-app feedback mechanisms *(Tullis and Albert 2024, pp. 100-150)* involves:

- Providing short surveys after chatbot interactions
- Offering easy-to-use bug reporting tools

### 3.3.4 Ethical Design Patterns

### 3.3.5 Dark Pattern Avoidance

Avoiding dark patterns *(Brignull 2023)* includes:

- Providing transparent information about chatbot capabilities
- Offering clear opt-out options for data collection

### 3.3.6 Attention Economy Awareness

Addressing attention economy concerns *(Williams 2024, pp. 10-30)* involves:

- Designing for focussed, purposeful interactions
- Avoiding addictive design patterns

### 3.3.7 Psychological Impact

The potential psychological effects of AI-based support on students must be carefully considered:

#### Avoiding Over-Reliance

Mitigating the risk of over-reliance on AI for emotional support *(Miner et al. 2022, p. 746)* involves:

- Clearly communicating AI's role as a supplement, not replacement, for human support
- Integrating the chatbot with pre-existing human counselling services and enabling direct connections to appropriate resources and non-AI support

## 3.4 Monitoring, Evaluation, and Continuous Improvement

### 3.4.1 AI and Machine Learning Architecture

#### Algorithm Selection

Applying machine learning algorithms *(Géron 2024, pp. 25-50)* includes:

- Rely upon supervised learning and human-in-the-loop models for classification tasks (e.g., identifying at-risk students). "Human-centric AI" is a key principle.
- Implementing reinforcement learning for adaptive responses

#### Natural Language Processing (NLP) Frameworks

Implementing NLP frameworks *(Jurafsky and Martin 2024, pp. 1-15)* involves:

- Utilising BERT or GPT-based models for understanding context and intent, allowing for the provision of specified and bespoke services.
- Custom training on domain-specific data for career advice, mental health support, and academic guidance.

### 3.4.2 Conversational Metrics

Measuring conversational metrics *(Quarteroni et al. 2024, pp. 1-32)* includes:

- Assessing response accuracy and relevance
- Tracking task completion rates

#### Telemetry Systems

Setting up telemetry systems *(Vadapalli 2023, pp. 30-60)* includes:

- Implementing real-time data collection on user interactions
- Ensuring privacy-preserving logging mechanisms

### 3.4.3 Usage Analytics

Implementing usage analytics *(Beasley 2023, pp. 50-100)* includes:

- Tracking common queries and pain points
- Analysing conversation flows and completion rates

#### Sentiment Analysis

Performing sentiment analysis *(Liu 2023, pp. 50-100)* includes:

- Analysing emotional tone of student interactions
  (N.B. It is worth noting that certain interpretations of EU legislation suggest that AI systems capable of deciphering emotions may be inherently considered unethical. However, given the rapid pace of technological innovation, particularly in the AI sector, it is plausible that the European Parliament may adopt a nuanced approach, interpreting the law's intent rather than adhering strictly to its literal wording) *(Dignum 2023, pp. 150-175)*
- Identifying potentially distressed students

### 3.4.4 Continuous Learning and Adaptation

Enhancing chatbot performance over time:

#### Research Collaboration

Fostering research collaboration *(Dillenbourg 2023, pp. 50-100)* involves:

- Establishing partnerships with universities for cutting-edge research
- Publishing findings to contribute to the broader field

# 4 Reflection and Recommendations

Key recommendations for the 'Alice' project:

- Establish a robust ethical framework and governance structure. Data security and data governance is of paramount importance.
- Implement state-of-the-art AI and machine learning technologies with a focus on inclusive outputs, free from negative bias.
- Train and prepare staff to work effectively alongside AI systems, not in lieu of them.
- Prioritise students, cater to diverse student needs.
- Develop comprehensive monitoring and evaluation systems for continuous improvement
- Ensure continued compliance with legislature.

# A Appendix A: Data Protection Impact Assessment Template

[Include a template or example of a Data Protection Impact Assessment (DPIA) tailored for the 'Alice' chatbot project]

# B Appendix B: Regulatory Standards

## DATA PROTECTION LEGISLATION

### General Data Protection Regulation (GDPR)

"

- Establishing a lawful basis for processing: Consent or legitimate interests
- Implementing data subject rights: Access, rectification, erasure, portability
- Appointing a Data Protection Officer (DPO)

"

*(European Union 2016)*

### UK Data Protection Act 2018

"

- Adhering to specific provisions for processing personal data in educational contexts
- Implementing safeguards for processing special category data (e.g., health information)

"

*(UK Government 2018)*:

### Children's Online Privacy Protection Act (COPPA)

"

- Obtaining parental consent for students under 13
- Implementing limited data collection and retention policies

"

*(Federal Trade Commission 2023)*

## EDUCATION SECTOR REGULATIONS

### Education and Skills Act 2008

"

- Fulfilling the duty to promote the well-being of students
- Implementing safeguarding responsibilities in digital environments

"

*(UK Government 2008)*

### Keeping Children Safe in Education

"

- Implementing online safety measures for educational technology
- Providing staff training on digital safeguarding

"

*(Department for Education 2024a)*

**Special Educational Needs and Disability (SEND) Code of Practice**

"

- Ensuring accessibility requirements for digital learning tools are met
- Considering personalised support for students with SEND

"

*(Department for Education 2024b)*

## PROFESSIONAL STANDARDS AND GUIDELINES

**BCS Code of Conduct**

"

- Considering public interest in development decisions
- Maintaining professional competence and integrity
- Fulfilling duty to relevant authorities

"

*(British Computer Society 2024, pp. 1-5)*

**ACM Code of Ethics and Professional Conduct**

"

- Contributing to society and human well-being
- Avoiding harm in system design and implementation
- Maintaining honesty and trustworthiness

"

*(Association for Computing Machinery 2024, pp. 1-4)*

**IEEE Ethically Aligned Design**

"

- Preserving human rights in AI systems
- Ensuring transparency and accountability in AI decision-making
- Implementing privacy-by-design principles

"

*(Institute of Electrical and Electronics Engineers 2024, pp. 2-5)*

## INDUSTRY-SPECIFIC STANDARDS

**ISO/IEC 27001:2022**

"

- Conducting risk assessment and management
- Implementing information security controls
- Establishing continuous improvement processes

"

*(International Organization for Standardization 2022)*

**Learning Tools Interoperability (LTI) Standards**

"

- Ensuring secure integration with existing learning management systems
- Enabling data portability and interoperability

"

*(IMS Global Learning Consortium 2024)*

**Web Content Accessibility Guidelines (WCAG) 2.2**

"

- Ensuring the chatbot interface is perceivable, operable, understandable, and robust
- Maintaining compatibility with assistive technologies

"

*(World Wide Web Consortium 2023)*

## ETHICAL AI FRAMEWORKS

**UNESCO Recommendation on the Ethics of Artificial Intelligence**

"

- Protecting human rights and fundamental freedoms
- Promoting diversity and inclusiveness in AI systems
- Ensuring transparency and explainability of AI decisions

"

*(UNESCO 2021)*

**OECD AI Principles**

"

- Ensuring AI benefits people and the planet
- Designing AI systems that respect the rule of law, human rights, democratic values, and diversity

"

*(Organisation for Economic Co-operation and Development 2023)*

**EU Ethics Guidelines for Trustworthy AI**

"

- Implementing human agency and oversight in AI systems
- Ensuring technical robustness and safety
- Maintaining privacy and data governance

"

*(European Commission 2024)*

## CONTINUOUS COMPLIANCE/ONGOING AUDITING IN PERPETUITY

### Regular Compliance Audits

"

- Performing annual data protection audits
- Engaging third-party security assessments

"

*(Information Commissioner's Office 2024)*

### Ethics Review Board Oversight

"

- Conducting periodic reviews of chatbot decisions and outcomes
- Establishing stakeholder feedback mechanisms

"

*(AI Ethics Board Best Practices 2024)*

### Continuous Professional Development

"

- Providing regular training on evolving legal and ethical standards
- Obtaining certification in AI ethics for key personnel

"

*(Chartered Institute of Personnel and Development 2024)*

# References


AI Ethics Board Best Practices (2024). *Guidelines for establishing and managing AI ethics boards*. URL: https://www.aiethicsboard.org/best-practices (visited on 04/10/2024).

AI Ethics Guidelines for Education (2024). *Best practices for implementing AI in educational settings*. URL: https://www.aieducationethics.org/guidelines (visited on 04/10/2024).

American Psychological Association (2024). *Guidelines for the practice of telepsychology*. URL: https://www.apa.org/practice/guidelines/telepsychology (visited on 04/10/2024).

Anastasiou, D. and R. Schäler (2023). *Translating Vital Information: Localisation, Internationalisation, and Globalisation*. Routledge.

Annus, N. (2023). "Chatbots in Education – the impact of Artificial Intelligence based ChatGPT on Teachers and Students". In: *International Journal of Advanced Natural Sciences and Engineering Researches* 7.4, pp. 366–370.

Armstrong, T. (2023). *Neurodiversity in the Classroom: Strength-Based Strategies to Help Students with Special Needs Succeed in School and Life*. ASCD.

Arrieta, A.B. et al. (2022). "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI". In: *Information Fusion* 58, pp. 82–115.

Article 29 Working Party (2018). *Guidelines on consent under Regulation 2016/679*. URL: https://ec.europa.eu/newsroom/article29/items/623051 (visited on 04/10/2024).

Association for Computing Machinery (2024). *ACM Code of Ethics and Professional Conduct*. URL: https://www.acm.org/binaries/content/assets/about/acm-code-of-ethics-and-professional-conduct.pdf (visited on 10/04/2024).

Association for Computing Machinery Conference on Fairness, Accountability, and Transparency (2024). "Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency". In: New York: ACM.

Barocas, S., M. Hardt, and A. Narayanan (2021). *Fairness and Machine Learning: Limitations and Opportunities*. URL: https://fairmlbook.org/ (visited on 04/10/2024).

Beasley, M. (2023). *Practical Web Analytics for User Experience: How Analytics Can Help You Understand Your Users*. 2nd ed. Morgan Kaufmann.

Bickmore, T.W. et al. (2021). "Patient and consumer safety risks when using conversational assistants for medical information: an observational study of Siri, Alexa, and Google Assistant". In: *Journal of Medical Internet Research* 20.9, e11510.

Blundell, Boyd (2020). *Computer Ethics and Professional Responsibility*. Oxford, UK: Oxford University Press.

Bradbury, A. (2024). *Successful Presentation Skills*. 6th ed. Kogan Page.

Brignull, H. (2023). *Dark Patterns: Inside the interfaces designed to trick you*. URL: https://www.darkpatterns.org/ (visited on 04/10/2024).

British Computer Society (2024). *Code of Conduct for BCS Members*. URL: https://www.bcs.org/membership-and-registrations/become-a-member/bcs-code-of-conduct/ (visited on 04/10/2024).

Chartered Institute of Personnel and Development (2024). *Continuing Professional Development: Guidelines and Best Practices*. CIPD. URL: https://www.cipd.co.uk/knowledge/fundamentals/people/development/continuing-professional-development-factsheet (visited on 04/10/2024).

Department for Education (2024a). *Keeping Children Safe in Education: Statutory guidance for schools and colleges*. URL: https://www.gov.uk/government/publications/keeping-children-safe-in-education--2 (visited on 04/10/2024).

— (2024b). *Special Educational Needs and Disability Code of Practice: 0 to 25 years*. URL: https://www.gov.uk/government/publications/send-code-of-practice-0-to-25 (visited on 04/10/2024).

Dignum, V. (2023). *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Cham, Switzerland: Springer Nature.

Dillenbourg, P. (2023). *Artificial Intelligence in Education: Promises and Challenges*. Cambridge, UK: Cambridge University Press.

El Emam, K. and L. Arbuckle (2023). *Anonymizing Health Data: Case Studies and Methods to Get You Started*. 2nd ed. O'Reilly Media.

European Commission (2024). *Ethics Guidelines for Trustworthy AI*. URL: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai (visited on 04/10/2024).

European Union (2016). *General Data Protection Regulation (GDPR)*. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj (visited on 04/10/2024).

Federal Trade Commission (2023). *Children's Online Privacy Protection Rule: A Six-Step Compliance Plan for Your Business*. URL: https://www.ftc.gov/tips-advice/business-center/guidance/childrens-online-privacy-protection-rule-six-step-compliance (visited on 04/10/2024).

Floridi, L. and J. Cowls (2023). "A Unified Framework of Five Principles for AI in Society". In: *Harvard Data Science Review* 1.1.

Géron, A. (2024). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 3rd ed. O'Reilly Media.

Goodman, E., M. Kuniavsky, and A. Moed (2023). *Observing the User Experience: A Practitioner's Guide to User Research*. 3rd ed. Morgan Kaufmann.

Harper, S. and Y. Yesilada (2024). *Web Accessibility: A Foundation for Research*. 3rd ed. Springer.

Hartzog, W. (2023). *Privacy's Blueprint: The Battle to Control the Design of New Technologies*. Harvard University Press.

Holtzblatt, K. and H. Beyer (2024). *Contextual Design: Design for Life*. 3rd ed. Morgan Kaufmann.

Hoober, S. and E. Berkman (2023). *Designing Mobile Interfaces: Patterns for Interaction Design*. 2nd ed. O'Reilly Media.

IMS Global Learning Consortium (2024). *Learning Tools Interoperability (LTI) Standards*. URL: https://www.imsglobal.org/activity/learning-tools-interoperability (visited on 04/10/2024).

Information Commissioner's Office (2024). *Data protection guidance*. URL: https://ico.org.uk/for-organisations/guide-to-data-protection/ (visited on 04/10/2024).

Institute of Electrical and Electronics Engineers (2023). *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*. URL: https://standards.ieee.org/industry-connections/ec/autonomous-systems/ (visited on 04/10/2024).

— (2024). *IEEE Code of Ethics*. Professional Guidelines. New York, NY: IEEE.

International Organization for Standardization (2022). *ISO/IEC 27001:2022 Information security management systems — Requirements*. International Standard ISO/IEC 27001:2022. Geneva, Switzerland: ISO.

Josuttis, N.M. (2023). *Cloud Native Transformation: Practical Patterns for Innovation*. Addison-Wesley Professional.

Jurafsky, D. and J.H. Martin (2024). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. 3rd ed. Upper Saddle River, NJ: Prentice Hall.

Liu, B. (2023). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. 2nd ed. Cambridge University Press.

Mehrabi, N. et al. (2023). "A survey on bias and fairness in machine learning". In: *ACM Computing Surveys* 54.6, pp. 1–35.

Miner, A.S. et al. (2022). "Key considerations for incorporating conversational AI in psychotherapy". In: *Frontiers in Psychiatry* 10, p. 746.

Organisation for Economic Co-operation and Development (2023). *OECD Principles on Artificial Intelligence*. URL: https://www.oecd.org/going-digital/ai/principles/ (visited on 04/10/2024).

Quarteroni, S. et al. (2024). "Chatbot Evaluation Metrics: State of the Art and Future Directions". In: *Dialogue & Discourse* 12.1, pp. 1–32.

Stallings, W. (2023). *Cryptography and Network Security: Principles and Practice*. 8th ed. Pearson.

Tullis, T. and B. Albert (2024). *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. 3rd ed. Morgan Kaufmann.

UK Government (2008). *Education and Skills Act 2008*. URL: https://www.legislation.gov.uk/ukpga/2008/25/contents (visited on 04/10/2024).

— (2018). *Data Protection Act 2018*. URL: https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted (visited on 04/10/2024).

UNESCO (2021). *Recommendation on the Ethics of Artificial Intelligence*. URL: https://en.unesco.org/artificial-intelligence/ethics (visited on 04/10/2024).

Vadapalli, S. (2023). *DevOps and Site Reliability Engineering (SRE) Handbook: Guide to Site Reliability Engineering and DevOps Practices*. Packt Publishing.

Vaughan, J.W. (2024). *Human-in-the-Loop Machine Learning: Active learning and annotation for human-centered AI*. Manning Publications.

Williams, J. (2024). *Stand Out of Our Light: Freedom and Resistance in the Attention Economy*. Cambridge University Press.

World Wide Web Consortium (2023). *Web Content Accessibility Guidelines (WCAG) 2.2*. URL: https://www.w3.org/TR/WCAG22/ (visited on 04/10/2024).

Yesilada, Y., G. Brajnik, and S. Harper (2023). "How much does web accessibility cost?" In: *Proceedings of the 20th International Conference on World Wide Web*, pp. 1–10.