ISO/IEC TR 24368:2022-08 (E)

Information technology - Artificial intelligence - Overview of ethical and societal concerns

Contents

Page

Foreword v

Introduction vi

1 Scope 1

2 Normative references 1

3 Terms and definitions 1

4 Overview 3
4.1 General 3
4.2 Fundamental sources 4
4.3 Ethical frameworks 6
4.3.1 General 6
4.3.2 Virtue ethics 6
4.3.3 Utilitarianism 6
4.3.4 Deontology 6

5 Human rights practices 7
5.1 General 7

6 Themes and principles 8
6.1 General 8
6.2 Description of key themes and associated principles 8
6.2.1 Accountability 8
6.2.2 Fairness and non-discrimination 9
6.2.3 Transparency and explainability 9
6.2.4 Professional responsibility 10
6.2.5 Promotion of human values 10
6.2.6 Privacy 11
6.2.7 Safety and security 11
6.2.8 Human control of technology 12
6.2.9 Community involvement and development 12
6.2.10 Human centred design 13
6.2.11 Respect for the rule of law 13
6.2.12 Respect for international norms of behaviour 13
6.2.13 Environmental sustainability 14
6.2.14 Labour practices 14

7 Examples of practices for building and using ethical and socially acceptable AI 15
7.1 Aligning internal process to AI principles 15
7.1.1 General 15
7.1.2 Defining ethical AI principles the organization can adopt 15
7.1.3 Defining applications the organization cannot pursue 15
7.1.4 Review process for new projects 15
7.1.5 Training in AI ethics 16
7.2 Considerations for ethical review framework 16
7.2.1 Identify an ethical issue 16
7.2.2 Get the facts 16
7.2.3 List and evaluate alternative actions 17

7.2.4 Make a decision and act on it 17
7.2.5 Act and reflect on the outcome 17

8 Considerations for building and using ethical and socially acceptable AI 17
8.1 General 17
8.2 Non-exhaustive list of ethical and societal considerations 17
8.2.1 General 17
8.2.2 International human rights 18
8.2.3 Accountability 18
8.2.4 Fairness and non-discrimination 18
8.2.5 Transparency and explainability 18
8.2.6 Professional responsibility 19
8.2.7 Promotion of human values 20
8.2.8 Privacy 20
8.2.9 Safety and security 20
8.2.10 Human control of technology 21
8.2.11 Community involvement and development 21
8.2.12 Human centred design 21
8.2.13 Respect for the rule of law 21
8.2.14 Respect for international norms of behaviour 22
8.2.15 Environmental sustainability 22
8.2.16 Labour practices 22

Annex A (informative) AI principles documents 23

Annex B (informative) Use case studies 33

Bibliography 42