ISO/IEC 23894:2023-02 (E)

Information technology - Artificial intelligence - Guidance on risk management

Contents

Page

Foreword iv

Introduction v

1 Scope 1

2 Normative references 1

3 Terms and definitions 1

4 Principles of AI risk management 1

5 Framework 5
5.1 General 5
5.2 Leadership and commitment 5
5.3 Integration 6
5.4 Design 6
5.4.1 Understanding the organization and its context 6
5.4.2 Articulating risk management commitment 8
5.4.3 Assigning organizational roles, authorities, responsibilities and accountabilities 8
5.4.4 Allocating resources 8
5.4.5 Establishing communication and consultation 8
5.5 Implementation 9
5.6 Evaluation 9
5.7 Improvement 9
5.7.1 Adapting 9
5.7.2 Continually improving 9

6 Risk management process 9
6.1 General 9
6.2 Communication and consultation 9
6.3 Scope, context and criteria 9
6.3.1 General 9
6.3.2 Defining the scope 10
6.3.3 External and internal context 10
6.3.4 Defining risk criteria 10
6.4 Risk assessment 11
6.4.1 General 11
6.4.2 Risk identification 11
6.4.3 Risk analysis 14
6.4.4 Risk evaluation 15
6.5 Risk treatment 15
6.5.1 General 15
6.5.2 Selection of risk treatment options 15
6.5.3 Preparing and implementing risk treatment plans 16
6.6 Monitoring and review 16
6.7 Recording and reporting 16

Annex A (informative) Objectives 18

Annex B (informative) Risk sources 21

Annex C (informative) Risk management and AI system life cycle 24

Bibliography 26