DIN SPEC 92001-3:2023-08 (E)

Artificial Intelligence - Life Cycle Processes and Quality Requirements - Part 3: Explainability; Text in English

Inhalt

Seite

Foreword 4

Introduction 6

1 Scope 7

2 Normative references 7

3 Terms and definitions 7

4 Life Cycle, Stakeholders, and Information Needs 9
4.1 General 9
4.2 AI System Life Cycle 9
4.3 Stakeholder Roles 9
4.4 Interactions & Information Needs 9
4.4.1 General 9
4.4.2 Develop 11
4.4.3 Provide 11
4.4.4 Use 11
4.4.5 Evaluate 12

5 Opacity 12
5.1 General 12
5.2 What is Opacity? 12
5.3 Sources of Opacity 12
5.3.1 General 12
5.3.2 Technical Illiteracy 13
5.3.3 Intentional Obfuscation 13
5.3.4 System Complexity 13
5.4 Effects of Opacity 13
5.4.1 Effects on Development 13
5.4.2 Effects on Provision 14
5.4.3 Effects on Use 14
5.4.4 Effects on Evaluation 15

6 Explainability 15
6.1 General 15
6.2 Explanation & Explainability 15
6.3 Explanans: Content and Form 15
6.4 Interpretation & Interpretability 16
6.5 When is explanation needed? 16
6.5.1 General 16
6.5.2 Reliability 17
6.5.3 Stakes 17
6.6 Risks of Explanation 17

7 Providing Explanations 18
7.1 General 18
7.2 Classification of Explanation Methods 18
7.3 Explanation Qualities 18
7.3.1 Explanation documentation 18
7.3.2 Fidelity and accuracy 19

7.3.3 Stability and consistency 19
7.3.4 Psychological appropriateness 19
7.3.5 Conceptual Alignment 20
7.3.6 Explanation Variety 20
7.3.7 Completeness 20
8 Explainability by Design 21
8.1 General 21
8.2 Requirement Analysis 21
8.3 Explanation Technical Design 21
8.4 Explanation Validation 22
Bibliography 23

Figures

Figure 1 — The Explainability by Design process 21

Tables

Table 1 — Stakeholder roles (columns) and life cycle stages (rows) 10