

# DIN SPEC 92001-2:2020-12 (E)

## Artificial Intelligence - Life Cycle Processes and Quality Requirements - Part 2: Robustness

---

### Contents

	Page
<b>Foreword .....</b>	<b>3</b>
<b>Introduction.....</b>	<b>5</b>
<b>1 Scope .....</b>	<b>6</b>
<b>1.1 Field of Application.....</b>	<b>6</b>
<b>1.2 Limitations .....</b>	<b>6</b>
<b>2 Normative References.....</b>	<b>7</b>
<b>3 Terms and Definitions .....</b>	<b>7</b>
<b>3.1 General Terminology .....</b>	<b>7</b>
<b>3.2 Terminology – Adversarial Robustness .....</b>	<b>7</b>
<b>3.3 Terminology – Corruption Robustness.....</b>	<b>9</b>
<b>3.3.1 Corruption Robustness – General terminology .....</b>	<b>9</b>
<b>3.3.2 Distributional Shift/ Dataset Shift.....</b>	<b>10</b>
<b>3.3.3 Sample Selection Bias.....</b>	<b>10</b>
<b>4 AI Quality Metamodel.....</b>	<b>11</b>
<b>5 Robustness.....</b>	<b>13</b>
<b>5.1 Introduction to AI Robustness.....</b>	<b>13</b>
<b>5.2 Requirements and Guidelines on Risk Management.....</b>	<b>14</b>
<b>5.2.1 Overview.....</b>	<b>14</b>
<b>5.2.2 Scope, Context, and Criteria.....</b>	<b>17</b>
<b>5.2.3 General Goals and Objectives.....</b>	<b>19</b>
<b>5.3 Requirements specific to Adversarial Robustness.....</b>	<b>20</b>
<b>5.3.1 Scope, Context, and Criteria.....</b>	<b>20</b>
<b>5.3.2 Threat Model Analysis .....</b>	<b>20</b>
<b>5.3.3 Likelihood &amp; Impact Analysis.....</b>	<b>22</b>
<b>5.3.4 Robustness Evaluation .....</b>	<b>27</b>
<b>5.3.5 Mitigations .....</b>	<b>29</b>
<b>5.4 Requirements specific to Corruption Robustness.....</b>	<b>32</b>
<b>5.4.1 Scope, Context, and Criteria.....</b>	<b>32</b>
<b>5.4.2 Threat Model Analysis .....</b>	<b>32</b>
<b>5.4.3 Likelihood &amp; Impact Analysis.....</b>	<b>34</b>
<b>5.4.4 Robustness Evaluation .....</b>	<b>37</b>
<b>5.4.5 Mitigations .....</b>	<b>39</b>
<b>6 Implementation Guidelines .....</b>	<b>41</b>
<b>Bibliography .....</b>	<b>43</b>