AI Cloud Service Compliance Criteria Catalogue (AIC4)

Bundesamt für Sicherheit in der Informationstechnik. (2023, May 22). AI Cloud Service Compliance Criteria Catalogue (AIC4). BSI. <https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/CloudComputing/AIC4/AI-Cloud-Service-Compliance-Criteria-Catalogue_AIC4.pdf?__blob=publicationFile&v=4>

The document is a guide for auditors to evaluate the compliance of an AI service provider.

Compliance Criteria are divided into 8 areas

|  |  |
| --- | --- |
| Preliminary Criteria | Ensure that the AI service provider demonstrates compliance of its  AI service with general cloud computing compliance criteria according to C5 (Cloud Computing Compliance Criteria Catalogue) as well as consistent documentation of policies and instructions applicable for the AI service |
| Security and robustness | Secure the AI service and improve its robustness against attacks by  ensuring confidentiality and integrity of data along the service’s  training pipeline through testing against malicious input data as well  as the implementation of countermeasures. |
| Performance and Functionality | Ensure accurate processing in line with performance requirements  for the AI service by training, evaluating and testing extensively before deployment. Tailored assessment methods are used in accordance with requirements for accurate processing. Training of AI model(s) within the scope of the AI service follows established procedures. |
| Reliability | Ensure continuous operation of the AI service in the production environment and investigate possible failures through appropriate procedures for resource management, logging, failure management and backups. |
| Data Quality | Ensure that data used by the AI service (e.g. training and test data)  comes from trustworthy sources, fulfills quality criteria, is annotated correctly and is protected adequately. |
| Data Management | Ensure that a framework providing guidelines for data handling, data quality, and data access as well as handling of data sources for the AI service is in place. |
| Explainability | Provide measures so that users can understand and explain decisions made by the service when necessary while depending on the sensitivity of the target application making the lack of explainability transparent to users. |
| Bias | Provide methods to assess and mitigate possible bias within the service or training data using appropriate techniques. |