NACHVOLLZIEHBARER ABLAUF

Von externen Signalen zu geprüften Entscheidungen

Ein sichtbares Signal kann eine Prüfung begründen, ohne bereits eine Schwachstelle, einen Schaden oder ein Kontrollversagen zu belegen. Halten Sie Evidenz, Begründung und Entscheidung miteinander verbunden.

> **Beobachten / Strukturieren / Korrelieren / Bewerten / Prüfen / Berichten**

01 Beobachten

Definieren Sie Dienst, Prüffrage und zulässigen Rahmen. Erfassen Sie, was sichtbar ist, wo es gefunden wurde und wann. Sichern Sie eine datierte Referenz oder einen geeigneten Evidenzauszug.

02 Strukturieren

Beschreiben Sie Asset, Quelle, Zeitstempel, relevanten Auszug und Einschränkungen einheitlich. Trennen Sie direkte Beobachtungen von Interpretationen und fehlenden Informationen.

03 Korrelieren

Vergleichen Sie verfügbare unabhängige Quellen. Prüfen Sie Inhaberschaft, Zeitpunkte und Kontext. Mehrere Kopien derselben Quelle liefern keine unabhängige Bestätigung.

04 Bewerten

Legen Sie dar, was die Evidenz stützt, welche plausible Folge denkbar ist und welche Voraussetzungen dafür nötig wären. Prüfen Sie alternative Erklärungen.

05 Prüfen

Eine prüfende Person kontrolliert Argumentation, Evidenzqualität und verbleibende Unsicherheit. Halten Sie abweichende Einschätzungen und Angaben fest, die die Bewertung ändern könnten.

06 Berichten

Nennen Sie Befund, praktische Bedeutung, nächsten Schritt, verantwortliche Stelle und Prüftermin. Bewahren Sie genügend Herkunftsdaten, damit Dritte die Argumentation nachvollziehen können.

> Beobachtet: direkt durch eine protokollierte Quelle belegt. Abgeleitet: eine Interpretation mit ausdrücklicher Begründung. Unbekannt: nicht festgestellte Information.
>
> Ein automatisierter Score ist ein Eingabewert für die Prüfung. Halten Sie ihn an seine Evidenz und Grenzen gebunden.

CHECKLISTE 01

Externe Exposition

Erfassen Sie für jeden Punkt Ja / Nein / Unbekannt / Nicht anwendbar sowie eine stützende Quelle und eine verantwortliche Person. Nutzen Sie die Checkliste als Prüfleitfaden, nicht als automatischen Punktestand.

- ☐ **01 Dienstinventar**
  Ist bekannt, welche Domains, Anwendungen und öffentlichen Schnittstellen zur Organisation gehören?
- ☐ **02 Inhaberschaft**
  Lassen sich eigene Dienste von Lieferanten, aufgegebenen Systemen und unbeteiligten Nachahmer-Domains unterscheiden?
- ☐ **03 Öffentliche Informationen**
  Legen veröffentlichte Seiten und Dokumente nur die Informationen offen, die die Organisation bewusst öffentlich bereitstellt?
- ☐ **04 Konto- und Zugangswege**
  Sind Registrierung, Anmeldung, Wiederherstellung und Supportwege klar und konsistent gestaltet?
- ☐ **05 Drittanbieter**
  Sind externe Dienste bekannt, die den öffentlichen Nutzerablauf unterstützen?
- ☐ **06 Datenschutzhinweise**
  Findet eine besuchende Person aktuelle Angaben zur Datenverarbeitung und zu Kontaktwegen?
- ☐ **07 Missbrauchsmeldung**
  Besteht ein funktionierender Meldeweg für Identitätsmissbrauch, verdächtige Nachrichten oder Sicherheitsvorfälle?
- ☐ **08 Wartungssignale**
  Sind öffentliche Sicherheitshinweise, Dokumentationen und Dienstinformationen auf aktuellem Stand?
- ☐ **09 Evidenzqualität**
  Wurden vermutete Auffälligkeiten gegen naheliegende harmlose Erklärungen abgeglichen?
- ☐ **10 Nachverfolgung**
  Verfügt jeder relevante Punkt über eine verantwortliche Person, einen nächsten Schritt und ein Prüfdatum?

> **Mit Kontext priorisieren**
> Was ist exponiert? Wer könnte betroffen sein? Welche zusätzlichen Bedingungen sind erforderlich? Wie belastbar ist die Evidenz? Welche Maßnahme verringert Unsicherheit oder Auswirkung?

Eine Versionsangabe, eine erreichbare Seite oder ein verdächtiger Domainname allein beweist weder eine Ausnutzung noch eine böswillige Inhaberschaft.

CHECKLISTE 02

KI in Bildung und Diensten für Kinder und Jugendliche

Ein praxisnaher Austausch zwischen Dienstverantwortlichen, pädagogischen Fachkräften und zuständigen Spezialisten. Erfassen Sie zu jedem Punkt eine Antwort, stützende Evidenz und offene Fragen.

01 Zweck. Welche Aufgabe soll die KI übernehmen, und welche Entscheidungen bleiben ausdrücklich bei Menschen?

02 Nutzende. Wer soll den Dienst nutzen, und wie spiegeln sich Schutz- und Altersbedürfnisse in der Gestaltung wider?

03 Daten. Welche Angaben gelangen in das System? Lässt sich die Aufgabe mit weniger sensiblen oder synthetischen Beispielen prüfen?

04 Lebenszyklus. Sind Speicherung, Aufbewahrungsfristen, Löschung und eventuelle Weitergabe klar erläutert?

05 Grenzen. Welche Themen, Handlungen oder Entscheidungen muss das System ablehnen oder an Menschen übergeben?

06 Verlässlichkeit. Wie werden unbelegte Antworten, Widersprüche und fehlende Evidenz erkannt und kenntlich gemacht?

07 Menschliche Aufsicht. Wer darf eine Ausgabe prüfen, korrigieren, unterbrechen oder anfechten?

08 Kommunikation. Erkennen Nutzende eindeutig, dass sie mit KI interagieren, und verstehen sie die jeweiligen Grenzen?

09 Zugänglichkeit. Verstehen die vorgesehenen Nutzenden alle Hinweise und können sie Melde- und Hilfswege leicht nutzen?

10 Erprobung. Spiegeln Prüffälle die vorgesehenen Aufgaben, Sprachen und absehbaren Fehlerbedingungen realistisch wider?

11 Vorfallreaktion. Wer nimmt Hinweise oder Beschwerden entgegen, und welche konkreten Schritte folgen nach einer Meldung?

12 Änderungskontrolle. Wer genehmigt wesentliche Änderungen, und wann muss der Dienst neu bewertet werden?

> **Vor einem Pilotbetrieb**
> Halten Sie zulässige Aufgaben, Testdaten, die verantwortliche Prüfperson, erforderliche Nachweise für den Start sowie Bedingungen zum Anhalten des Pilotbetriebs schriftlich fest.

FALLBEISPIEL

Wenn die Evidenz noch nicht ausreicht

Vollständig fiktives Beispiel. Es wird kein realer Dienst, keine reale Organisation und kein realer Vorfall dargestellt.

> SYNTHETISCHER FALL / Eine kleinere Bildungseinrichtung prüft den Einsatz eines KI-Assistenten für Verwaltungsfragen.
>
> Prüffrage: Liegt ausreichende Evidenz vor, um einen begrenzten Pilotbetrieb mit Angaben zu Lernenden zuzulassen?

Evidenz	Dokumentierte Beobachtung	Einschränkung
E-01	Die bereitgestellte Produktbeschreibung besagt, dass eingereichte Fragen gespeichert werden können.	Eine Beschreibung belegt noch keine tatsächliche Systemkonfiguration.
E-02	Die vorgelegten Angaben nennen keine Speicherfrist für eingereichte Anfragen.	Schutz- und Fristregeln können bestehen, wurden aber nicht vorgelegt.
E-03	Das geplante Pilotkonzept sieht auch Testfragen mit Angaben zu Lernenden vor.	Der Pilotbetrieb hat noch nicht begonnen.

Interpretation

Identifizierbare Lernendendaten könnten ein vermeidbares Datenumgangsrisiko erzeugen, solange Speicherfristen und Löschverhalten ungeklärt bleiben.

Unbekannte Punkte

Tatsächlich konfigurierte Speicherfristen, reales Löschverhalten, Zugriffe des Anbieters sowie die Möglichkeit, Speicherung vollständig zu deaktivieren.

Entscheidung

HOLD für identifizierbare Lernendendaten: Nutzen Sie während der Erstevaluierung ausschließlich synthetische Fragen. Halten Sie identifizierbare Lernendendaten aus dem Pilotbetrieb heraus, bis der Dienstverantwortliche ausreichende Evidenz für eine geprüfte Entscheidung vorlegt.

Nächste Evidenz und Prüfanlass

Fordern Sie eine aktuelle Beschreibung der Aufbewahrungs- und Löschkontrollen, die einschlägige Konfiguration sowie einen dokumentierten Prüfnachweis an. Bewerten Sie den Fall neu, sobald Nachweise eingehen oder sich Pilotzweck beziehungsweise Datenbestand ändern.

Die Schlussfolgerung betrifft die Ausreichendheit der Evidenz für diesen geplanten Einsatz. Sie stellt keinen Sicherheitsvorfall und keinen Rechtsverstoß fest.

Quellenherkunft und vertiefende Hinweise: Abgeleitet aus der veröffentlichten Methodik von SKOPION, dem Prüffeld für kinder- und jugendnahe KI sowie der IGF-Standbeschreibung. Dies sind reine Lehrmaterialien ohne Benchmark-Zusagen.

skopion.de/de/methodik

skopion.de/de/igf-2026

Offizielles IGF-Profil / Village-Stand #20 SKOPION Intelligence

DRUCKVORLAGE

Evidenz- und Entscheidungsprotokoll

Verwenden Sie pro Prüffrage ein eigenes Protokollblatt. Fügen Sie weitere Nachweise oder Anlagen bei, falls der vorgesehene Platz nicht ausreicht.

Protokoll-ID / Bewertungsdatum / Prüfende Person

Prüffrage / Dienst oder Asset / Prüfrahmengrenzen und Ausschlüsse

ID	Quelle / Datum	Beobachtung	Einschränkung

Interpretation / alternative Erklärungen / unbekannte Punkte oder Widersprüche

Plausible Folge und erforderliche Voraussetzungen

Vertrauensniveau (Niedrig / Mittel / Hoch) und Begründung

Anmerkungen oder abweichende Einschätzung der menschlichen Prüfung

Entscheidung (inkl. HOLD) und Begründung

Maßnahme / verantwortliche Stelle / Zieldatum

Erforderliche Evidenz zur Klärung oder Neubewertung / Prüfdatum und Auslöser

Ergebnis nach menschlicher Prüfung