

AI Intelligence Review

Et evidensbaseret review af AI-systemers styrker, friktion og risici — med hvert fund sporbart

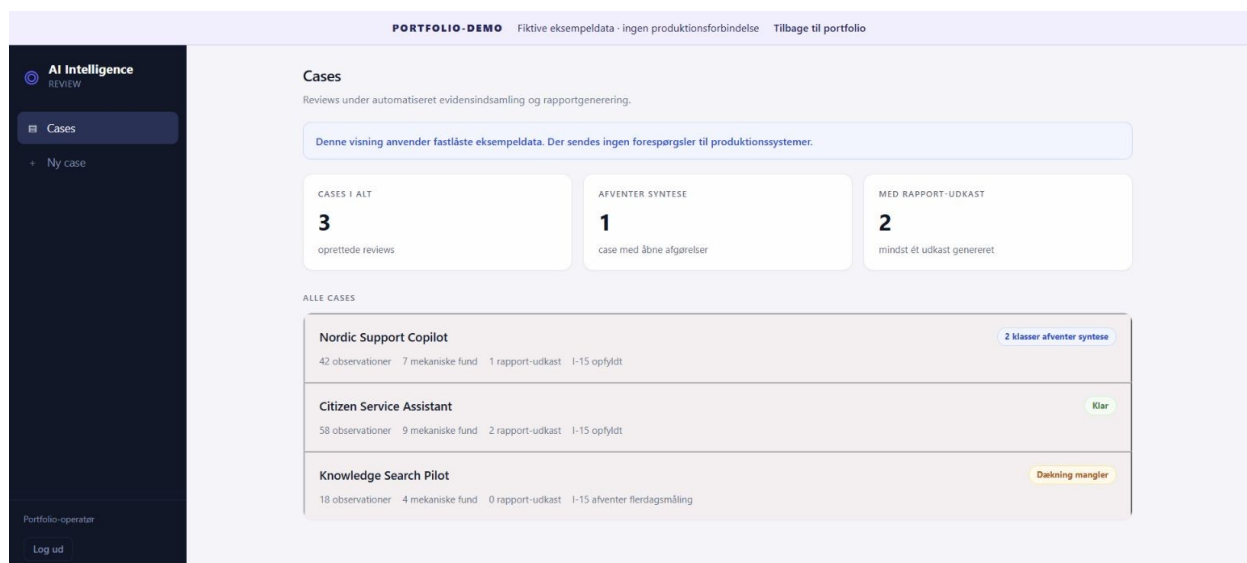
Produktiseret AI-review · 15 fundklasser · låst testbatteri · mekanisk evidens + menneskelig syntese

Hvad AI Intelligence Review er

AI Intelligence Review (AIR) er et **produktiseret, uafhængigt review** af AI-systemer. Scope, metode og rapportstruktur er faste, så vurderingen kan gentages og forklares. Produktet undersøger ikke, om en AI virker imponerende i en enkelt samtale; det undersøger, om adfærden er dokumenterbar, konsekvent og forsvarlig over tid.

Metoden arbejder med **15 fundklasser**. Syv kan opdares mekanisk, mens otte kræver menneskelig eller kontrolleret AI-assisteret syntese. Den arbejdsdeling er bevidst: maskinen samler og strukturerer evidensen, men den må ikke opfinde alvor, konklusion eller sikkerhed, hvor grundlaget ikke bærer det.

Portfolio-versionen viser den virkelige operatørflade med **fiktive, fastlåste eksempeldata**. Den opretter ingen rigtige cases og sender ingen forespørgsler til produktionssystemer. Dermed kan arbejdsgangen undersøges sikkert uden adgang til kunde-, login- eller driftsdata.



Caseoverblikket viser målegrundlag, syntesestatus og rapportberedskab uden at skjule uafklarede trin.

Hovedkomponenter

Fra låst afgrænsning til sporbar rapport. Hver komponent er beskrevet ved hvad den **gør**, med et konkret eksempel.

Låst scope og testbatteri — Case, mål, kanaler og spørgsmål fastlåses før første måling. Ændringer kræver en ny version i stedet for en stille omskrivning. *Eksempel: Et spørgsmål kan ikke tilføjes bagefter blot fordi de første svar peger i en interessant retning.*

Evidensindsamling — Svar, DOM-uddrag, målinger og udvalgte screenshots registreres med tidspunkt, kanal og reference. *Eksempel: Et fund om kildeklarhed peger tilbage på den konkrete passage og den observation, der bærer påstanden.*

15-klasses domænemodel — Alle relevante fund placeres i en fast taksonomi, så dækning og fravær kan kontrolleres på tværs af reviews. *Eksempel: Systemet kan skelne mellem kildeklarhed, uensartet afvisning, robusthed og driftsstabilitet uden at blande mekanismerne sammen.*

Mekaniske detektorer — Syv klasser findes deterministisk ud fra observationsloggen og det låste facit. *Eksempel: Gentagne svar kan sammenlignes maskinelt, så uensartet adfærd opdages uden at være afhængig af reviewerens hukommelse.*

Syntesegate — De otte vurderingskrævende klasser stoppes ved en eksplicit menneskelig afgørelse; systemet må ikke fabricere severity. *Eksempel: Hvis konsekvensen ikke kan begrundes, forbliver casen markeret som 'afventer syntese'.*

Konsekvensbaseret prioritering — Fund sorteres efter dokumenteret bruger- og driftskonsekvens frem for et kunstigt totalscorekort. *Eksempel: Et alvorligt, veldokumenteret enkeltfund kan stå øverst, selv om resten af systemet fungerer godt.*

Rapportassembly — Fund, evidens, falsifikatorer, styrker og anbefalinger samles i en fast rapportstruktur med ét kalibreret verdikt. *Eksempel: En anbefaling kan spores baglæns fra rapporten til fundet, observationen og den oprindelige evidens.*

Operatørfladen — Cases, status, mekaniske fund, syntese og rapportudkast samles i én rolig arbejdsflade. *Eksempel: Operatøren kan se, at I-15's flerdagskrav er opfyldt, før rapporten kan betragtes som klar.*

The screenshot displays the 'AI Intelligence REVIEW' interface. On the left is a dark sidebar with navigation options: 'Cases', 'Ny case', and 'Portfolio-operatør'. The main area shows a case titled 'Nordic Support Copilot' with details like 'Case-id: nordic-support-copilot' and 'Låst: 08-08-2026'. Below this are four summary cards: 'OBSERVATIONER' (42), 'MEKANISKE FUND' (7), 'SYNTSE-KLASSE' (2), and 'RAPPORT-UDKAST' (1). A green status bar indicates 'I-15 (opetid/drift): 2,3 dages målespænd — opfylder det automatiserede ≥2-dages-krav.' Below this is a table of 'MEKANISKE FUND' with three entries: 'AIR-13-004 — Kildeklarhed' (HØJ), 'AIR-19-002 — Uensartet afvisning' (MELLEM), and 'AIR-111-001 — Robusthed' (LAV). Each entry includes a brief description and a severity indicator.

Casevisningen forbinder observationer, mekaniske fund, syntese og rapportudkast i samme sporbare arbejdsgang.

Arbejdsgangen

- **1. Afgræns og lås.** Reviewets mål, adgangsvej, kanaler, facit og testbatteri skrives ned før målingen starter.
- **2. Mål og registrér.** Den samme rejse gennemføres på tværs af relevante variationer og tidspunkter; observationerne gemmes append-only.
- **3. Detektér og syntetiser.** Mekaniske klasser findes automatisk. De resterende kræver en begrundet afgørelse med påstand, evidens og falsifikator.
- **4. Prioritér og rapportér.** Værste dokumenterede konsekvens først, men med samme beviskrav til styrker. Rapporten viser både hvad der holder, og hvad der skal ændres.
- **5. Kvalitetssikr frisk.** En frisk læser eller frisk kontekst kontrollerer, at rapporten kan forstås uden arbejdsdokumentet, og at ingen konklusion er stærkere end evidensen.

PORTFOLIO-DEMO Fiktive eksempeldata · ingen produktionsforbindelse Tilbage til portfolio

AI Intelligence
REVIEW

Cases

+ Ny case

Ny case

Definér reviewets mål, scope og det låste testbatteri.

Portfolio-demoen viser arbejdsgangen, men opretter eller kontakter ingen eksterne systemer.

CASE

Navn
Eksempel på AI-assistent

Systemejer
Demoorganisation

REVIEW-SCOPE

Mål
https://example.invalid/api
Produktionsforbindelser er slået fra i portfolio-demoen.

Opret og lås case Annullér

En ny case begynder med et eksplicit mål og en låst afgrænsning — ikke med efterrationaliserede spørgsmål.

Hvor jeg ser det gå hen

Kernen er på plads; næste fase handler om at gøre den kontrollerede arbejdsdeling endnu stærkere:

- **Færdiggør det assisterede syntesetrin.** LLM eller menneske kan foreslå alvor og formulering, men kun gennem en synlig gate og med den ansvarlige reviewer som ejer af konklusionen.
- **Flere lag af samme sandhed.** Generér ledelsesresumé, fuld rapport og følgebrev fra den samme master, så korte formater aldrig udvikler deres egen version af fakta.
- **Stærkere kalibrering.** Brug afsluttede reviews til at forbedre spørgsmål, falsifikatorer og kvalitetsbar — versioneret og uden at omskrive historiske resultater.
- **Fra operatørværktøj til skalerbar praksis.** Adskil reviewer, frisk læser og kundeansvarlig, mens metode, evidenskrav og leveringsgates forbliver uændrede.

Afslutning

AI Intelligence Review gør AI-kvalitet til en **sporbar undersøgelse frem for en mavefornemmelse**. Produktet kombinerer software-engineering, metode- og evidensdisciplin, menneskelig dømmekraft og en rapportmotor, der nægter at udfylde huller med selvsikker tekst. Som showcase demonstrerer det både den tekniske pipeline og den sværere produktbeslutning: at automatisere det gentagelige uden at automatisere ansvaret væk.