LEGAL DOCUMENT

AI, human review and accountability

Explains AI assistance, human review and accountability.

VERSION	0.6.1
REVISED (UTC)	2026-09-26 01:50:50 UTC
BODY SHA-256	8cafda6d203b0b33d66c9009a475f354865e38a8559776a50a456db05e4de930
PUBLISHED AT	https://www.onsiren.com/legal/ai/editions/0.6.1

This policy explains the permitted uses of AI, accountability, human review and the safeguards required for each connected service.

01

AI scope and accountability

AI may support screening, routine case administration, request triage and sourced explanations. OnSiren's team administers the current Legal Centre and handles escalation. Accountable people own consequential legal interpretation, exceptions, disputed findings, sensitive disclosures and remediation under the approved policy. No additional staffing or salary commitment is implied.

Significant decisions with no meaningful human involvement need an Article 22A–C assessment. Such decisions cannot rely wholly or partly on Article 6(1)(ea). Special-category significant automated decisions carry additional restrictions. Routine administrative automation stays within its approved purpose and safeguards.

02

The public question helper and tracked requests

The initial public question is matched to current published Legal Centre documents within this service. An explicit follow-up can send the current and previous question and selected published passages to OpenAI for an explanation. The server keeps the API key outside the browser, uses store:false, enables no tools or web research and bounds the request by time, input and output limits. The setting does not establish zero provider retention. Questions can contain personal data; provider terms, retention, locations and transfer arrangements require assessment.

Ordinary question and answer text is not saved in application request or conversation records. Hashed-owner usage metadata is retained to enforce limits. The helper retrieves published documents only, with no private-case retrieval, restricted research, unpublished documents, face search, live product-record search or autonomous sensitive-disclosure connection. Missing evidence, unavailable providers and unsupported

questions receive document guidance or a route to human review. An answer does not establish legal verification or a completed action.

A person explicitly submits the Request Centre form to create a tracked request and receipt. Ordinary helper questions are not automatically recorded for human triage. Automatic recognition and referral of an actual rights request or complaint typed into the general helper is not implemented. The substance of an actual submission remains decisive, and accountable people must preserve its original receipt and arrange handling through any available channel. The question explanation and request-submission steps must be clear before information is entered.

03

Continuous checks and limits

The operational worker checks this module's deadlines, fixed agreement integrity, review dates, document status and known integration gaps while the service runs. Bounded legal-source checks compare retrieved source content and flag changes for review. Each result must identify scope, evidence, last successful check, errors and remaining limitations. An unresolved legal change remains pending through a fetch failure.

These checks cover their recorded systems and rules. Missing production feeds, failed checks and unassessed law remain visible. A successful comparison supplies no guarantee that all relevant legal developments have been found or interpreted. The displayed schedule must match the current implementation.

04

Future production AI must use approved providers, minimised inputs, defined roles, retention, transfer safeguards, evaluations, cost limits and audit records. No general model training on officer alerts, privacy cases or union casework is authorised by this policy.

Before personal data reaches a supplier, verify its role, documented instructions, confidentiality, security, sub-processors, permitted use, retention and locations. Assess every restricted transfer and apply its mechanism and required safeguards. Cost controls and current provider settings must be evidenced separately.

05

Alert moderation and challenge

The agreed alert design uses AI screening and UK-based human review of identifiable third-party material before release. Reviewers must have the relevant information, training and authority to change or reject a recommendation and escalate uncertainty. Reporter names remain outside routine moderation access; exceptional identity access requires separate authority and logging.

The intended 24/7 human coverage and technical detection gates require operational evidence. Human challenge and correction routes must remain available. The agreed policy specifies safeguards and coverage intent; it does not establish a particular employment model, new hire or full-time post.