DfE generative AI product safety self-assessment

Sendix · Against the DfE standards updated 19 January 2026 · self-assessed · based on DPA v1.4 and DPIA v0.7 · November 2026

Date	15 October 2026 (pen test status updated 26 October 2026)
Product	Sendix web app, run by Sendix AI Ltd (company SC883770, registered in Scotland). School SEND staff upload Education, Health and Care Plans (England) and Child's Plans and IEPs (Scotland). Sendix flags vague provision, shows the basis for each flag, and suggests wording. A member of staff accepts, edits or rejects every suggestion.
Users	Adult school staff only (SENCOs, SEND leads, IT leads, billing admins). Pupils and parents do not use Sendix.
DfE page	Generative AI: product safety standards, last updated **19 January 2026**. DfE first published this as "product safety expectations" (22 January 2025). The January 2026 update renamed it and added sections on cognitive development, emotional and social development, mental health and manipulation.
Assessed by	Sendix AI Ltd, self-assessed; not independently verified

Summary

We checked Sendix against all 63 points in the DfE standards, grouped as DfE groups them. **15 meet**, **12 partly meet**, **1 does not meet**, and **35 are not applicable**. Most of the "not applicable" points are in sections DfE marks as relevant to learner-facing products (filtering, monitoring and reporting, cognitive development, emotional and social development, mental health). Pupils never use Sendix, so these do not apply, but we say what Sendix does in each case. The main gaps are: we have not published an accuracy evaluation; there is no documented testing of the current AI model before it was adopted; the independent penetration test has not yet been done (booked for January 2027; until then no more than 3 organisations use real plans, each with its DPO's written acceptance); Cyber Essentials is applied for but not certified; and there is no formal complaints procedure for safety issues.

The rule for real data. Sendix processes a school's real plans only on its instructions, under a Data Processing Agreement (DPA) the school signs and Sendix countersigns, and only after Sendix's separate release approval. Sendix's independent penetration test is not yet done (booked for January 2027). Until it is complete, Sendix releases real pupil data to no more than 3 organisations, each only with a current Cyber Essentials certificate, cross-school tests passing on the version in use, and its data protection officer's written acceptance that the test is still to come.

How the AI is used. Sendix sends parts of a plan, with the pupil's name and other identifiers removed, to one AI model: Qwen3-235B (`qwen.qwen3-235b-a22b-2507-v1:0`) on Amazon Bedrock in London (eu-west-2). The model is called directly in London. Sendix blocks any model request that does not match a current, signed approval record for a UK-only, no-retention route. Vague wording is found by fixed rules, not by the AI. The AI explains the flag using legal passages Sendix supplies, and proposes wording. By default a proposed fix must be made of sentences copied exactly from the plan or the evidence. If there is not enough support, Sendix offers no fix and leaves it to staff.

Stated purpose

Expectation	Status	What Sendix does	Evidence

Clearly state the product's intended purpose and use cases	Meets	Sendix helps SEND staff review support plans. It flags vague or unsupported provision, cites the basis and suggests wording. Staff decide every change. It makes no decision about a child.	`docs/compliance/dpia/DPIA_v0.7.md` (Step 1, 2.3); security pack (`backend/app/routers/onboarding.py`, `_SECURITY_PACK`); this document
Give information on target demographic and learning focus	Meets	Users are adult school staff. The plans are about children and young people with SEND (0 to 25 in England). Focus: the quality of provision in EHCPs, Child's Plans and IEPs. Sendix is not a learning tool.	`DPIA_v0.7.md` (2.4 Scope); `docs/compliance/dpia/DPIA_addition_scotland_plans.md`
Review the intended purpose when features change	Meets	New features that use plan data are off by default and need a DPIA addition first. Examples: termly follow-through and the guided IEP review. The DPIA is reviewed on any new data, supplier, model or incident.	`backend/app/core/config.py` (`FOLLOW_THROUGH_ENABLED`, `IEP_REVIEW_ENABLED`); `docs/compliance/dpia/DPIA_addition_follow_through_DRAFT.md`; `DPIA_v0.7.md` (Step 7)
Do not exaggerate impact or capabilities; back claims with robust, transparent evidence	Partly meets	We audit our own claims against the code, and an automated test stops the security pack claiming things that are not built. The app tells staff that suggestions can be wrong. But we have not published an accuracy evaluation, so we cannot yet show evidence for how accurate the flags and suggestions are.	`docs/business-claims-technical-audit-2026-09-18.md`; `docs/validation/business-claims-release-2026-09-18.md`; `backend/tests/test_security_pack_claims.py`; `DPIA_v0.7.md` (R6)

Educational use cases

Expectation	Status	What Sendix does	Evidence
State which DfE use cases apply	Meets	**Administrative and management**: it supports staff with plan review and compliance with statutory guidance (the SEND Code of Practice in England; Education Scotland and GIRFEC guidance in Scotland). None of the learner-facing use cases apply.	`DPIA_v0.7.md` (Step 1); this document
Provide sufficient evidence that the product meets its stated purpose and these standards	Partly meets	This self-assessment, the DPIA, the DPA and the security pack set out what the product does. There is no published accuracy evaluation, and no independent review of this assessment.	`DPIA_v0.7.md`; `docs/compliance/dpa/Sendix_DPA_v1.4.md`; `_SECURITY_PACK`

Filtering

DfE marks this section as relevant to learner-facing products. Pupils do not use Sendix, so each point is not applicable. What still applies: Sendix has no free-text chat. Staff cannot ask the model open questions. The model only sees plan extracts, legal passages Sendix supplies and, where AI analysis in review preparation is on (off for real pupil data), filtered review evidence, and its suggestions are shown to staff, never to pupils. Sendix does not let users share content with each other and does not search live websites.

Expectation	Status	What Sendix does	Evidence

Users reliably prevented from generating or accessing harmful or inappropriate content	Not applicable	Staff-only. No open prompt box. Suggestions are built from sentences already in the plan or evidence (default policy), so the model cannot introduce unrelated content into a suggestion.	`backend/app/core/config.py` (`CONVERSE_GROUNDING_POLICY = "extractive"`); `backend/app/services/converse/remediation/cascade.py`
Filtering maintained throughout a conversation	Not applicable	There are no conversations. Each model call is a single, fixed task.	`backend/app/services/converse/`
Filtering adjusted by risk, age and need (including SEND)	Not applicable	No pupil users.	—
Multimodal moderation (languages, images, misspellings)	Not applicable	Sendix reads text from PDF and Word plans only. It does not generate images.	`backend/app/services/document_parser.py`
Full moderation on any device, including BYOD	Not applicable	No pupil users. Staff access is the same web app on any device.	—
Moderation based on the context of the conversation	Not applicable	No conversations.	—
Filtering updated for new types of harmful content	Not applicable	No pupil users.	—

Monitoring and reporting

DfE marks this section as relevant to learner-facing products. Pupils do not use Sendix, so each point is not applicable. Staff activity is logged (first row).

Expectation	Status	What Sendix does	Evidence
Robust activity logging: inputs and responses, performance metrics	Not applicable	For staff use, Sendix stores each AI output with the finding it belongs to, and records who accepted, edited or rejected it and when. Security events are kept in an audit log that cannot be edited. Model call counts, errors and response times are measured without recording content.	`backend/app/models/unified_findings.py` (`FindingTransition.model_output`, `actor`; `Diagnosis.model_version`); `backend/app/models/audit_log.py`; `backend/app/services/model_policy.py` (call metrics); `docs/compliance/retention-and-deletion.md`
Alert local supervisors to searches for, or access to, harmful content	Not applicable	No pupil users and no search.	—
Alert and signpost the user to support when prohibited content is attempted	Not applicable	No pupil users.	—
Real-time, age-appropriate notice when content is blocked	Not applicable	No pupil users.	—

Sendix AI Ltd · SC883770 · registered in Scotland · ICO ZC264381 · privacy@sendix.ai

Alert supervisors to disclosures that may be a safeguarding issue	Not applicable	Pupils do not type into Sendix. Plans can contain sensitive information already held by the school. Sendix does not scan plans for safeguarding concerns. Those remain with the school's own processes.	—
Hold and confirm the Designated Safeguarding Lead's contact details for alerts	Not applicable	Sendix does not send safeguarding alerts, so it does not collect DSL details.	—
Reports and trends on access to prohibited content, readable by non-experts	Not applicable	No pupil users.	—
Report to teachers on cognitive offloading, emotional engagement and each learner's usage time	Not applicable	No pupil users.	—

Security

DfE marks this section as relevant to learner and teacher-facing products, so it applies.

Expectation	Status	What Sendix does	Evidence
Robust protection against jailbreaking to reach prohibited material	Partly meets	There is no chat, so staff cannot give the model instructions directly. Text inside a plan could still try to instruct the model. Defences: fixed tasks and output formats; Scottish plan prompts tell the model to treat text as data; suggestions must be exact sentences from the sources; a suggestion is dropped if Sendix would flag its own wording; nothing is used unless staff accept it. Missing: we have no recorded adversarial (red-team) testing, and the prompt-injection filter only covers the older drafting feature, which is not reachable.	`backend/app/services/converse/remediation/cascade.py` (`_fails_own_checks`); `backend/app/services/converse/detection/childs_plan.py`; `backend/app/services/converse/remediation/iep.py`; `backend/app/services/input_sanitizer.py`
Prevent unauthorised changes that could reprogram the product	Meets	Code changes go through GitHub, which requires two-factor sign-in for our organisation. Changes pass required checks and a staging deploy before production. Secrets sit in Azure Key Vault. The model can only be changed by updating a signed approval record; a mismatch blocks every request. These controls have not been independently tested.	`docs/compliance/patch-policy.md`; `docs/compliance/cyber-essentials-readiness.md`; `docs/model-hosting.md`; `backend/app/services/model_policy.py`; `backend/tests/test_model_policy.py`

Administrators can set different permission levels	Meets	Roles: SENCO, SEND lead, IT lead, billing admin, trust SEND lead. Only SENCOs and SEND leads at the school (and the trust's SEND lead) can see plans. IT leads change settings but cannot see plan content. Billing admins see billing only. Each role can only be invited by set roles. Sendix administrator accounts cannot read plan content in the application.	`backend/app/core/plan_access.py`; `backend/app/routers/onboarding.py` (`_INVITER_ROLES`); `backend/tests/test_cross_tenant_access.py`; `docs/compliance/admin-access.md`
Regular bug fixes and updates, promptly applied	Meets	High and critical security updates are applied within 14 days. Dependencies, container images and code history are scanned on every change and weekly.	`docs/compliance/patch-policy.md`; `.github/workflows/security.yml`; `.github/dependabot.yml`
Test new versions or models for safety before release	Partly meets	Every release runs the automated test suite (over 100 test files), including cross-school access tests, and is deployed to staging and health-checked first. Missing: the model-change procedure has no evaluation step, and there is no recorded evaluation of the current model's output quality before it was adopted.	`docs/compliance/patch-policy.md`; `docs/model-hosting.md` ("Changing model"); `backend/tests/`; `backend/scripts/converse_validate.py`
Robust password protection or authentication	Meets	Passwords of at least 12 characters, checked against a deny-list. Lockout after 5 failed attempts. Authenticator-app MFA is mandatory for SENCOs, SEND leads and Sendix admins; optional for other school roles. Sessions end after 30 minutes idle. Single sign-on is not available.	`docs/compliance/admin-access.md`; `backend/app/services/mfa.py`; `backend/app/core/passwords.py`; `backend/tests/test_mfa.py`
Compatible with the Cyber Security Standards for Schools and Colleges	Partly meets	In place: prompt patching, MFA on accounts that can see plans, access limited by role. Missing: Cyber Essentials was applied for on 9 October 2026 and the assessment is under way; we are not certified. The independent penetration test is not yet done (booked for January 2027; until it is, no more than 3 organisations use real plans.	`docs/compliance/cyber-essentials-readiness.md`; `docs/compliance/patch-policy.md`; `DPIA_v0.5.md` (R1, 26 Oct 2026)

Privacy and data protection

DfE marks this section as relevant to learner and teacher-facing products, so it applies. The school is the data controller; Sendix is its processor.

Expectation	Status	What Sendix does	Evidence
Clear privacy notice, shown at regular intervals, in age-appropriate language, covering data types, where data is processed and the legal basis	Partly meets	Staff see the privacy notice and terms at sign-up, and the accepted terms version is recorded. We give schools a family privacy notice insert explaining AI-assisted review, with a short pupils' version for Scotland. There is no pupils' version of the England insert. Sendix does not re-show the notice at set intervals. The public notice is maintained in our website code, not in this repository.	`docs/compliance/dpa/family-privacy-notice-insert.md`; `docs/compliance/dpa/family-privacy-notice-insert-scotland.md`; `docs/compliance/privacy-and-terms-release.md`; `backend/tests/test_terms_version.py`

DPIA during development and for the whole life of the tool	Meets	DPIA v0.7 approved by Sendix on 6 November 2026 (v0.6: 29 October 2026), with additions for Scotland and follow-through. Its consultation step is outstanding and the processor assessment is provisional. Reviewed every 6 months and on any new data, supplier, model or incident. The school's own sections are completed by the school, with a pre-filled pack for its DPO.	`docs/compliance/dpia/DPIA_v0.7.md`; `docs/compliance/dpia/DPIA_addition_scotland_plans.md`; `docs/compliance/dpo-pack/`
Let all parties meet controller and processor duties without overburdening each other	Meets	Our DPA v1.4 sets out each party's duties, including breach notice within 24 hours. Schools export their own data. Sendix can erase or restrict a pupil's plans on the school's written instruction within 10 working days. Departing schools get a deletion certificate.	`docs/compliance/dpa/Sendix_DPA_v1.4.md`; `docs/compliance/dsr-runbook.md`; `docs/compliance/retention-and-deletion.md`
Comply with data protection law and ICO codes, including the Children's code	Partly meets	Registered with the ICO (ZC264381). Data stays in the UK (Azure UK South; model in AWS London). Pupil names, NHS numbers, UPNs, postcodes and most written dates are removed from plan text before storage and before the model. Other people's names are removed on a best-effort basis. Health and social care sections are discarded on upload. Automatic deletion runs daily. Missing: we have not done a formal assessment against the ICO Children's code (Sendix is not used by children), the school's DPIA sections are not yet completed by any school, and none of this has been independently verified.	`DPIA_v0.7.md`; `docs/compliance/sub-processors.md`; `backend/app/services/third_party_names.py`; `backend/app/services/section_presence.py`; `docs/compliance/retention-and-deletion.md`
No commercial use of personal data, including model training, without a lawful basis	Meets	The DPA forbids using school data, or anything derived from it, to train or improve AI models, for cross-customer analytics or for marketing. The model provider does not store inputs or outputs or use them for training. The app loads no third-party analytics.	`Sendix_DPA_v1.4.md` (clause 2.4); `DPIA_v0.7.md` (R5, R11); `docs/model-hosting.md`; `backend/app/services/model_policy.py`

Intellectual property

DfE marks this section as relevant to learner and teacher-facing products, so it applies.

Expectation	Status	What Sendix does	Evidence
Do not use learners', teachers' or copyright owners' work for commercial purposes such as training or fine-tuning without consent	Meets	Uploaded plans and staff edits are never used to train or fine-tune models.	`Sendix_DPA_v1.4.md` (clause 2.4); `DPIA_v0.7.md` (2.2)

Do not collect, store or share inputs for commercial purposes, including product improvement and development, without permission	Meets	Plans are stored only to provide the service the school asked for, on the school's instructions. The school confirms it is entitled to upload each plan. Plans are not used for product development, benchmarking or analytics. They are deleted on the school's retention schedule.	`Sendix_DPA_v1.4.md` (clauses 2.4 and 2.5, Schedule 1); `docs/compliance/retention-and-deletion.md`

Design and testing

DfE marks this section as relevant to learner and teacher-facing products, so it applies.

Expectation	Status	What Sendix does	Evidence
For child-facing products: risk mitigations, child-centred design, testing with children	Not applicable	Sendix is not child-facing.	—
Sufficient testing with a diverse, realistic range of users and use cases	Partly meets	Tested on fictional English EHCPs and Scottish Child's Plans and IEPs, including a stress-test plan built to contain many known problems. Sendix staff test the live product as an ordinary SENCO with fictional plans. We have spoken to SENCOs about the product. Missing: there are no recorded structured user tests, and all testing so far has used fictional plans.	`docs/internal-testing.md`; `backend/data/pilot/pilot_ehcp_stress_test.docx`; `docs/validation/scotland-demo-review-2026-09-18.md`; `DPIA_v0.7.md` (Step 3)
Test new versions or models for safety before release	Partly meets	Same as under Security: automated tests and staging on every release, but no recorded evaluation of the current model before adoption.	`docs/compliance/patch-policy.md`; `docs/model-hosting.md`
The product consistently performs as intended	Partly meets	Vague wording is found by fixed rules, with automated tests. The AI must cite only passages Sendix supplies and must say when they do not support a conclusion. Each explanation is labelled "Retrieved basis for human assessment (not a verified legal conclusion)". A notice says AI suggestions can be wrong and are not legal advice. Only accepted changes reach the exported plan. Missing: no published accuracy evaluation (false positives, false negatives, citation accuracy).	`backend/app/services/provision_quality.py`; `backend/app/services/converse/detection/weasel.py`; `backend/app/services/converse/diagnosis/prompts.py`; `backend/app/services/converse/diagnosis/diagnoser.py`; `frontend/src/components/audit/ConversePanel.tsx`; `DPIA_v0.7.md` (2.3, R6)

Governance

DfE marks this section as relevant to learner and teacher-facing products, so it applies.

Expectation	Status	What Sendix does	Evidence
A clear risk assessment for the product	Meets	The DPIA has a risk register covering security, access, re-identification, model data use, wrong AI suggestions, retention and transfers, each with controls and an owner. Schools also get a residual-risk form for their DPO.	`DPIA_v0.7.md` (Steps 5 and 6); `docs/compliance/dpo-pack/residual-risk-acceptance.md`

Sendix AI Ltd · SC883770 · registered in Scotland · ICO ZC264381 · privacy@sendix.ai

A formal complaints mechanism, with quick escalation and resolution of safety issues	Does not meet	Schools can contact privacy@sendix.ai and hello@sendix.ai, and the DPA sets a 24-hour deadline for telling a school about a personal data breach. There is no written complaints procedure with steps and timescales for raising and resolving safety issues with the software.	`Sendix_DPA_v1.4.md` (clause 7); `DPIA_v0.7.md` (processor contact)
Policies and processes for AI safety decisions are made available	Partly meets	The rules exist and are written down: the model approval policy (UK-only, no retention, signed, expiring), the requirement that a person decides every change, and the default that suggestions must be copied from sources. A summary is published at sendix.ai/trust (the sub-processor list, with the model approval policy, and this assessment), and schools see more in the security pack and DPIA. The full internal documents are shared with schools on request rather than published.	`docs/model-hosting.md`; `backend/app/services/model_policy.py`; `DPIA_v0.7.md` (2.3); `_SECURITY_PACK`

Cognitive development

DfE marks this section as relevant to learner-facing products. Pupils do not use Sendix, so each point is not applicable. Sendix does suggest finished wording, but to professionals, who must decide on each suggestion.

Expectation	Status	What Sendix does	Evidence
Regular engagement with educators, child-safety experts, AI ethicists and psychologists	Not applicable	No pupil users. We engage with SENCOs about the product. We have no record of engagement with child-safety experts, AI ethicists or psychologists.	`DPIA_v0.7.md` (Step 3)
Child-safety training for technical teams	Not applicable	No pupil users. We have no record of child-safety training for our technical team.	—
Ongoing expert monitoring of the impact on learners' development	Not applicable	No pupil users.	—
Publish records of expert oversight and a child-development impact plan	Not applicable	No pupil users.	—
No final answers by default: progressive hints, ask for learner input first, full solution only after an attempt, friction or teacher approval between modes	Not applicable	No pupil users.	—
Track, detect and report cognitive offloading	Not applicable	No pupil users.	—

Emotional and social development

DfE marks this section as relevant to learner-facing products. Pupils do not use Sendix, so each point is not applicable. We describe what Sendix does where it is relevant to staff.

Expectation	Status	What Sendix does	Evidence
Do not anthropomorphise: function-based phrasing, no "I" statements, no names, avatars or characters, no claims of agency	Not applicable	No pupil users. In practice Sendix has no named assistant, avatar or chat. The panel is titled "AI review — diagnosis & suggested fixes" and outputs are short explanations and proposed wording.	`frontend/src/components/audit/ConversePanel.tsx`
Do not undermine real-world support networks or isolate the learner	Not applicable	No pupil users.	—
Do not prompt personal or emotional topics or cultivate relationships	Not applicable	No pupil users and no conversation.	—

Remind users that AI cannot replace real human relationships	Not applicable	No pupil users.	—
Default time limits: break prompts, hard limits, teacher override, "Stop for now" warning	Not applicable	No pupil users.	—
Do not artificially extend engagement	Not applicable	No pupil users. See also Manipulation below.	—
Record session durations and report each learner's level and nature of engagement	Not applicable	No pupil users.	—
Monitor emotional disclosures and concerning patterns, and notify the DSL	Not applicable	No pupil users.	—
Remember learner inputs only when needed; store minimum data; restrict access	Not applicable	No pupil users. For plan data, see Privacy and data protection.	—

Mental health

DfE marks this section as relevant to learner-facing products. Pupils do not use Sendix, so each point is not applicable.

Expectation	Status	What Sendix does	Evidence
Detect signs of learner distress	Not applicable	No pupil users.	—
Follow a tiered pathway when distress is detected (signposting, safeguarding flag)	Not applicable	No pupil users.	—
Use safe, supportive, non-validating language that directs learners to human help	Not applicable	No pupil users.	—
Safeguarding governance: child mental health expertise, child-safety training, a published crisis protocol	Not applicable	No pupil users. We have none of these in place.	—

Manipulation

DfE marks this section as relevant to learner and teacher-facing products, so it applies.

Expectation	Status	What Sendix does	Evidence
No manipulative or persuasive strategies: sycophancy, deception, unjustified confidence, social pressure, guilt or fear, threats, inappropriate rewards	Partly meets	The AI is told never to invent citations or specifics, and to state low confidence when the passages do not support a conclusion. Each explanation shows its confidence and sources. When there is not enough support, Sendix offers no fix rather than a plausible guess. For Scottish plans, which are not statutory, automated tests block legal-obligation language. There are no badges, rewards or peer comparisons. Missing: confidence labels come from the model and have not been checked against measured accuracy, and we have not tested outputs for sycophancy or overconfidence.	`backend/app/services/converse/diagnosis/prompts.py`; `backend/app/services/converse/remediation/prompts.py`; `backend/tests/test_non_statutory_language.py`; `frontend/src/components/audit/ConversePanel.tsx`

Sendix AI Ltd · SC883770 · registered in Scotland · ICO ZC264381 · privacy@sendix.ai

Do not exploit users: no designs that prolong use, steer to paid options, mix in advertising, or use dark patterns	Meets	The app has no adverts, no in-app upselling and no third-party analytics. Emails are transactional only (verification, invitations, password resets, plan-transfer notices). Accept, edit and reject are shown side by side, and nothing is applied without a decision.	`backend/app/services/email_service.py`; `frontend/src/components/audit/ConversePanel.tsx`; `frontend/package.json`

Sendix AI Ltd, company SC883770, registered in Scotland. ICO registration ZC264381. Contact: privacy@sendix.ai.