Intel e-Medix — Clinical & Technical Dossier

What the system does, how reports are produced, how accuracy must be understood, and what safeguards apply. Prepared for radiologists, ophthalmologists, cardiologists, physicians, hospitals, diagnostic centres, partners, compliance reviewers and technical reviewers. Version 1.0 — soft launch.

Status of this document. Every statement is marked either IMPLEMENTED (present in the running system today) or PLANNED (recommended, not yet built). Intel e-Medix has **no completed clinical validation study**, no published sensitivity/specificity/AUC figures, and no regulatory clearance (FDA, CE, DRAP or equivalent). No such figures appear anywhere in this document because none exist.

1. Executive explanation

Intel e-Medix is a software-assisted clinical reporting and health-literacy system. It accepts a medical document or medical image, extracts and organises the clinical content, and produces two distinct outputs: a structured clinician-facing report and a plain-English patient explanation. It is an assistive reporting layer — it does not replace interpretation by a qualified clinician, and final clinical responsibility always remains with the treating physician.

The architecture is not a single model. It combines deterministic logic (file validation, category heuristics, storage, retention, audit), one or more vision/language models accessed through an external inference gateway with a fallback provider chain, a fixed output schema that constrains what a report may contain, and application-level checks and cross-checks. Different steps are performed by different components.

```
Medical input
 -> file validation (type, size, integrity)            [IMPLEMENTED]
 -> modality / category identification + cross-check    [IMPLEMENTED]
 -> source-adequacy assessment                          [PARTIAL - narrative only]
 -> feature / finding extraction into fixed schema      [IMPLEMENTED]
 -> localisation, characterisation, severity grading    [IMPLEMENTED - text-level]
 -> differential consideration                          [IMPLEMENTED - narrative]
 -> impression / synthesis                              [IMPLEMENTED]
 -> triage level (routine / soon / urgent)              [IMPLEMENTED]
 -> recommendations + specialist referral               [IMPLEMENTED]
 -> ICD / procedure coding                              [PLANNED - not built]
 -> automated consistency & safety gate                 [PLANNED - not built]
 -> structured report + patient explanation + PDF       [IMPLEMENTED]
 -> physician review / sign-off                         [PROCESS REQUIREMENT, not enforced in
software]
 -> 48-hour automatic deletion                          [IMPLEMENTED]
```

2. What Intel e-Medix actually analyses

IMPLEMENTED Accepted inputs are PDF documents, JPEG/PNG/WebP images and MP4/MOV video. The system does **not** accept DICOM (`.dcm`) files or complete multi-slice CT/MRI studies, and it does not read PACS. It therefore cannot independently inspect a raw imaging volume to detect, measure or stage a primary pathology. Where the input is a written report, the system analyses the reporting radiologist's or pathologist's own text. Where the input is a single photographic image (for example a skin lesion or a fundus photograph), a vision model describes visible features — this is descriptive assistance, not a validated image-diagnostic function.

IMPLEMENTED The modality is never assumed from the filename alone. A text/filename heuristic and a vision classifier are run and compared across twelve categories (X-ray, CT, MRI, ultrasound, pathology/lab, skin, hair, cardiology, ophthalmology and related). Agreement raises the confidence label; disagreement lowers it to *low* and the interface asks the uploader to confirm the category before analysis proceeds.

PLANNED Explicit machine-readable determination of anatomical region and laterality as separate validated fields; DICOM ingestion, viewing and metadata extraction.

3. Source-quality control

Principle enforced in the reporting instructions: the system must not compensate for inadequate source quality by increasing diagnostic certainty.

IMPLEMENTED The clinician-facing report contains a mandatory *Limitations* section covering technical limitations, motion artefact, missing data and any factor affecting confidence, with an explicit "None — fully diagnostic" option. Where the source is unreadable or incomplete the report is instructed to state this and recommend repeat imaging or clinical reassessment rather than produce an interpretation.

PLANNED A quantitative pre-analysis quality gate scoring resolution, focus, exposure, contrast, field of view, artefact, obstruction, anatomical coverage and completeness, with a hard block on interpretation below threshold. Today this assessment is narrative and model-generated, not measured.

4. Finding detection and evidence traceability

IMPLEMENTED Findings are not free text. Every finding is emitted into a fixed schema with required fields: precise clinical term, measurement with unit and reference range where present, plain-English meaning, clinical significance, severity (`normal / low / moderate / high / unknown`), an **evidence quote taken verbatim from the source document**, and a supporting guideline reference. A finding that cannot carry a source quote is not supported by the schema's intent.

IMPLEMENTED Reporting instructions require calibrated language and forbid asserting certainty the evidence does not support. Accepted certainty ladder: *diagnostic of / highly suggestive of / consistent with / suggestive of / possible / indeterminate / cannot be determined from available data*.

PLANNED A machine-enforced OBSERVED / INFERRED / SUSPECTED / NOT ASSESSABLE flag on each finding. Today the distinction is expressed in wording, not as a validated structured field.

5. Reasoning, severity and triage

Dimension	How it is handled today	Status
Impression / synthesis	Structured clinician layout: indication, technique, comparison, findings, impression.	IMPLEMENTED
Disease vs. severity	Severity is a separate graded field per finding and for the study overall, distinct from the impression.	IMPLEMENTED
Differential	Alternatives discussed narratively when findings are non-specific.	IMPLEMENTED
Triage	Exactly one level assigned with a one-sentence justification, mirrored into a structured field (routine / soon / urgent) that drives dashboard prioritisation and emergency guidance.	IMPLEMENTED
Complications / further testing	Recommendations separate additional testing, specialist referral (with urgency and reason), follow-up and clinical correlation.	IMPLEMENTED
ICD / procedure coding	Not generated. No coding module exists.	PLANNED

Triage is derived from the clinical significance of the findings, not from a keyword match, and not every abnormal finding is escalated. Where emergency status depends on symptoms that are not present in the uploaded material, the report states that this cannot be assessed from the input.

6. Modality-specific reasoning

IMPLEMENTED Category detection selects a specialty-appropriate reporting template, so ophthalmic material is reported in ophthalmic terminology and radiology material in modality-specific radiological terms. The classifier is cross-checked (section 2) precisely to prevent a generic template being applied under the wrong specialty label. A low-confidence classification is surfaced for human confirmation rather than silently accepted.

PLANNED A hard modality-validation gate that refuses disease-specific interpretation logic until modality, region and laterality are confirmed.

7. Clinical references

IMPLEMENTED Reports cite recognised frameworks (ACR Appropriateness Criteria, Fleischner, BI-RADS/LI-RADS/TI-RADS, ADA Standards of Care, AHA/ACC, NICE, KDIGO, WHO) to explain the classification framework used.

Known limitation, stated plainly. Citations are produced by a language model and the system does not currently retrieve or verify the cited source. A citation must therefore be treated as an indication of the framework applied, never as proof that the conclusion is correct. Reviewing clinicians should verify any citation they intend to rely on. PLANNED : a curated, retrieval-backed reference store so that only verifiable citations can be emitted.

8. Report generation and output

IMPLEMENTED Two outputs are produced from one analysis. The clinician report follows structured-reporting convention (patient/report information, examination and modality, indication, technique, source information, comparison, findings, impression, severity, recommendations, limitations, review status). A separate compact clinical PDF targets one to three pages, dense and prioritised, with key findings, action plan, limitations and a physician-review notice. The patient explanation is kept separate and covers summary, findings in plain English, likely symptoms, recovery guidance, specialists, lifestyle guidance and references.

IMPLEMENTED Every report carries the boundary statement that the system organises and explains an existing medical report, is not a diagnosis, and does not replace a physician.

9. How uncertainty is handled

The system is instructed never to appear more capable by being more certain, and never to manufacture missing clinical information. Where evidence is insufficient the report states so directly — for example, that a feature cannot be definitively characterised on the available examination, that laterality could not be reliably established, or that OCT or clinical correlation is recommended. Missing fields are rendered as "Not provided", "Not assessable" or "Indeterminate" rather than filled by assumption.

10. Accuracy — how it must be understood

Intel e-Medix publishes no accuracy figure. No sensitivity, specificity, PPV, NPV, F1, AUC or overall-accuracy value has been established for this system, because no clinical validation study has been completed. Any such number quoted about Intel e-Medix today would be unsupported.

When validation is performed, performance must be reported multidimensionally, not as one headline number: sensitivity, specificity, positive and negative predictive value, accuracy, F1 and AUC/ROC where the task is classification, calibration, false-positive and false-negative rates, subgroup performance, inter-reader agreement, report completeness, and diagnostic agreement with qualified specialists.

Aggregate accuracy is misleading on its own: a system can score highly overall while failing on rare but clinically critical conditions. Evaluation must therefore report clinically significant failure modes separately from average performance.

Reference standard for future validation PLANNED

- Ground truth from independent expert interpretation, multi-specialist consensus, histopathology, confirmatory imaging, laboratory confirmation, established clinical diagnosis, or longitudinal follow-up as appropriate to the question.
- Cases used for validation must not have been used to tune prompts, templates or thresholds.
- Independent test set plus external, multi-centre validation across devices, populations, image qualities and deliberately ambiguous cases.

Physician–system agreement PLANNED

The most meaningful near-term quality measure is agreement with qualified clinicians, scored separately at the level of individual findings, disease classification, severity, laterality, urgency, impression and (once built) coding. Every disagreement should be classified by type and root cause and fed back into template revision.

11. Error and safety analysis

The following error classes must be monitored, with false negatives and clinically dangerous errors given priority: missed critical findings, false positives, over- and under-diagnosis, incorrect severity, incorrect laterality, incorrect modality, incorrect coding, unsupported certainty and internally contradictory statements.

IMPLEMENTED Today this is a manual process: outputs are sampled and reviewed against the source document, and physician and patient feedback is logged and folded into template revisions. Every report is traceable to the uploaded file and the uploading account through the audit trail. PLANNED: a structured error register with rate tracking per error class and per modality.

12. Internal report quality score

PLANNED — not yet implemented. The intended control is an automated pre-release scoring pass across diagnostic accuracy, finding accuracy, completeness, evidence support, clinical reasoning, severity, laterality, modality correctness, coding, recommendation quality, triage correctness, uncertainty calibration, internal consistency, safety and conciseness, with a target of ≥9.5/10 on every applicable dimension before release and automatic flagging for human review below that threshold.

> **Interpretation of the 9.5/10 target.** It is an internal writing- and quality-control threshold. It is **not** a diagnostic-accuracy claim and must never be presented to users as "95% accurate". A report that cannot reach the threshold on evidence must be flagged, not padded with invented content.

13. Physician review and clinical responsibility

```
system analysis -> structured report -> quality & safety checks -> physician review -> final
clinical report
```

IMPLEMENTED Physician-review notices are carried on report outputs and PDFs. PLANNED A software-enforced review queue with named reviewer, sign-off state and locked post-signature versioning. Until that exists, review is an organisational process obligation, not a system control. System output does not constitute a physician diagnosis under any configuration.

14. Illustrative example — retinal fundus photograph

Illustrative only; not a real patient case and not a diagnosis. Given a fundus photograph with features suggestive of diabetic retinopathy, correct system behaviour is: classify the modality as retinal/fundus imaging (never as a cardiology or other specialty label); describe the observable features; assess retinopathy severity and macular findings separately; state whether neovascularisation can or cannot be identified on the available image; decline to assert centre-involving macular oedema, because centre involvement requires OCT; recommend OCT and

ophthalmological assessment; and — once coding exists — emit a code matching the actual diagnosis, severity, laterality and macular-oedema status, or flag it for coding review. Where image quality prevents assessment, the correct output is an explicit limitation, not a graded severity.

15. Where the value lies

Intel e-Medix's contribution is consistency and communication, not autonomous diagnosis: a fixed structured schema instead of free text, evidence quoted from the source, modality-aware templates, explicit limitations and calibrated uncertainty, separated clinician and patient outputs, triage prioritisation, an audit trail linking every report to its uploader, and automatic 48-hour deletion of source files and analyses. No claim is made that the system outperforms clinicians or other systems; no comparative study has been run.

16. Final accuracy statement

> Intel e-Medix is designed to maximise the clinical quality, consistency and safety of software-assisted medical interpretation and report generation. Its actual diagnostic performance must be established through appropriately designed clinical validation against qualified reference standards. No single accuracy percentage should be used to represent performance across all diseases, modalities, populations and clinical scenarios.