

MedSeal-MBT 项目申报书

MoonBit DICOM 医学影像元数据清理工具

参赛账号	001-Elsa	项目性质	原创实现
项目方向	医学影像格式解析与安全工具	许可证	Apache-2.0
GitHub	https://github.com/001-Elsa/MedSeal-MBT	版本	0.1.1
Mooncakes	https://mooncakes.io/docs/001-Elsa/medseal	后端	Native

项目简介与场景

MedSeal-MBT 是使用 MoonBit 编写的轻量级 DICOM Part 10 文件解析、结构校验和基础元数据脱敏工具，面向开发测试、教学演示与数据交付前的直接身份字段清理。第一版仅接受 Explicit VR Little Endian；遇到其他 Transfer Syntax 会明确拒绝，不猜测解析医疗文件。

核心功能

- 解析 File Meta、常用文本/整数 VR、定义及常见未定义长度 Sequence。
- 检测截断、越界长度、非法 VR/UI/文本值、损坏 Sequence 与基础图像元数据问题。
- 替换或清空患者、机构、医生等直接身份字段，递归处理 Sequence，可选删除 Private Tag。
- Pixel Data 与 UID 原样保留；JSON 审计只记录 tag、字段名和动作，不记录原始值。
- 重新计算长度并写入临时文件，解析复检成功后原子替换；提供 inspect、validate、anonymize CLI。

技术路线与安全边界

项目按 byteio、model、parser、validator、deid、writer、app 分包。所有读取先检查剩余字节，并限制文件 256 MiB、单元素 128 MiB、元素 100,000 个、Sequence 深度 32 层。项目不解码压缩图像，不检查 Pixel Data 中烧录文字，不提供诊断，也不宣称完全匿名化或医疗法规合规。

交付与质量

交付公开 GitHub/GitLink 仓库、Apache-2.0 许可证、合成 DICOM 样本、可运行 Native CLI、审计示例、三平台 CI 和核心回归测试。Mooncakes 包 001-Elsa/medseal@0.1.1 已发布并构建成功。所有样本由仓库生成器创建，不含真实患者数据；直接依赖 moonbitlang/async 与 moonbitlang/x 均为 Apache-2.0。

联系方式：766903826@qq.com | A simple DICOM metadata validator and de-identification tool written in MoonBit.