医疗卫生机构患者个人敏感信息安全使用探究*

蔡雨蒙 伊向华 单红伟 王忠民 刘 云

（南京医科大学第一附属医院 南京 210029）

〔摘要〕 **目的/意义** 基于医疗卫生机构个人敏感信息使用典型场景，探究个人敏感信息使用安全实现方式。**方法/过程** 针对用户注册、内部利用与交互、医疗器械/可穿戴设备采集和信息披露等使用场景，分析个人敏感信息泄漏风险，探究身份鉴别、访问控制、数据加密、数据脱敏和检测审计等多种技术手段的适用场景及优缺点。**结果/结论** 技术手段的应用可以辅助医疗卫生机构进一步保障患者个人信息安全。

〔关键词〕 医疗卫生机构；个人敏感信息；使用安全

〔中图分类号〕R－058 〔文献标识码〕A 〔DOI〕10.3969/j.issn.1673－6036.2024.01.014

Study on the Safe Use of Sensitive Personal Information of Patients in Medical and Health Institutions

CAI Yumeng, YI Xianghua, SHAN Hongwei, WANG Zhongmin, LIU Yun

The First Affiliated Hospital with Nanjing Medical University, Nanjing 210029, China

〔**Abstract**〕 **Purpose/Significance** Based on the typical use scenario of sensitive personal information in medical and health institutions, the implementation of the secure use of sensitive personal information is explored. **Method/Process** For user registration, internal utilization and interaction, medical device/wearable device collection and information disclosure and other scenarios, the paper analyzes the risk of sensitive personal information leakage, explores the application scenarios, advantages and disadvantages of various technical means such as identity authentication, access control, data encryption, data desensitization and detection audit. **Result/Conclusion** The application of technical means can assist medical and health institutions to further protect the security of patients' personal information.

〔**Keywords**〕 medical and health institutions; sensitive personal information; use security

1 引言

2018 年 5 月欧盟《通用数据保护条例》正式施行，将健康信息、生物特征列入个人敏感信息范围并要求对个人敏感信息的处理必须合规。自此全球各国加速推进个人信息保护立法。2016—2020 年中国也相继出台相关法律法规，个人信息保护有了完善的法律依据。2016 年颁布的《中华人民共和国网络安全法》[1] 对个人信息收集与处理者提出合规的

〔修回日期〕 2023－06－12

〔作者简介〕 蔡雨蒙，工程师，发表论文 3 篇；通信作者：刘云，教授，主任医师。

〔基金项目〕 江苏省科技厅重点研发计划（项目编号：BE2020721）；江苏省南京市生命健康科技专项（项目编号：202205053）。

总体要求。2020 年《中华人民共和国民法典》[2]生效，对个人信息与隐私权作出明确定义。2021 年《中华人民共和国数据安全法》[3]与《中华人民共和国个人信息保护法》[4]生效，对个人信息、个人敏感信息进行了定义并且提出了个人主体可以主张的权利，以及个人信息保护的要求。

医疗卫生行业收集大量个人敏感信息，包括个人基本信息、生物特征、基因信息、健康信息等。用户就诊时没有选择余地，只能提供个人敏感信息如姓名、手机号、银行卡/医保卡号给医疗机构，且这些信息长期存留于医疗机构。由于医疗卫生行业数据价值很高，涉及国家生物安全、科研成果等，易遭外部入侵窃密或内部泄密。因此，本研究主要探讨针对医疗卫生机构个人敏感信息的访问与使用，如何采用多重手段保障个人信息安全。

2 医疗卫生机构个人敏感信息使用典型场景与风险

2.1 用户注册

用户注册使用时，会将个人敏感数据上传或增删改，且可能使用微信公众号、手机应用程序（App）挂号或下载诊疗报告等。典型注册－使用－下载场景如检验检查，用户在医院微信公众号或App 注册上传个人敏感信息，下载检验检查报告。在该场景下，面临的个人敏感信息安全风险为非最小化收集（过度收集非相关个人信息）、传输明文（在旧系统中可能传输不加密）、个人关键信息展示没有脱敏手段等。

2.2 内部利用与交互

医院内部利用场景以医疗机构内部诊疗信息采集、医师诊断、科室联合会诊为代表。交互使用的典型场景可以总结为3 类。一是跨机构交互，如跨地域或联合诊疗时个人敏感信息交换共享。二是大数据汇聚场景，如健康医疗大数据中心、医院科研部门等健康管理信息、大数据分析利用等；医保与医疗卫生机构信息交互用于结算理赔；医疗器械或软件的开发与维护等。三是医学合作研究/教学，该场景还可能涉及跨国研究、个人敏感信息出境传输等。

内部利用与交互使用面临的风险有以下几类。一是访问控制措施不足导致的非授权访问、扩大范围访问。二是个人信息未加密/脱敏导致的信息泄密。三是数据出境传输面临的合规风险[5]。

2.3 医疗器械/可穿戴设备采集

在医疗器械与可穿戴设备采集的场景下，以手环健康监测为例，监测到的脉搏、血氧含量等信息都归类于个人敏感信息，应当给予重点保护，在采集与使用过程中采取合适的安全措施，如访问控制、对个人敏感信息进行去标识/匿名化处理，操作审计等。中国信息通信研究院于2022 年7 月发布的《医疗物联网安全研究报告（2021 年）》[6]指出：医疗物联网设备激增，远程运维、漏洞无人关注等都导致了患者敏感数据泄漏风险大增。

2.4 信息披露

信息披露分为小范围公开（如科研成果内部共享）和公开披露（如流调信息），披露的内容以及个人敏感信息的处理，都要根据场景将个人信息进行加密、假名化/匿名化处理。有些高度敏感的信息，在数据共享前要经过相关监管机构审批（如涉及人类族谱、基因排序等科研信息）。

3 医疗卫生机构个人敏感信息使用安全探究

3.1 身份鉴别与访问控制相结合

身份鉴别与访问控制策略是个人敏感信息保护的必备基础之一。此外还有更复杂但是安全强度更高的零信任模式。这些手段组合使用，广泛适用于上述典型场景数据安全保护。

3.1.1 *身份鉴别* 在基础访问控制措施基础上，重点建设账号管理系统，形成基于角色访问授权的体系。为访问者授予账号，通过账号口令，结合多因素认证环节，确认访问者身份，进一步保障账号安全。多因素认证可采用短信验证码（如患者注册App 场景）、身份证/医保卡识别（如患者就诊缴费等环节）、电子证书（如医院内部人员访问场景）。

最终实现“访问者身份－账号与权限分配－可访问的数据资源”的映射关系。

3.1.2 访问控制措施 在医疗卫生机构内部，通常结合以下几类访问控制措施。（1）基于防火墙五元组的访问控制。实现对医疗机构内部的个人终端/科室地址段的访问控制或针对外部访问添加白名单，但是该措施的访问控制粒度较粗，不足以细粒度地保护个人敏感信息的访问和使用。（2）基于页面的访问控制。适用于浏览器/服务端（browser/server，B/S）架构的访问控制或通过域名间调用数据，该措施的颗粒度进一步细化为针对页面/应用的访问控制。（3）基于时间维度的访问控制。例如可设定医疗教学系统夜间禁止访问。（4）基于身份角色授权的访问控制。例如针对数据库角色授权，可以精细到针对库、表、列（列加密场景）级别。还可以结合证书授权（certificate authority，CA）、短信验证等手段，用于进一步确认身份，可以通过角色，对个人与应用访问授权进行绑定。（5）基于行为模型的访问控制。用户和实体行为分析技术可以针对医疗业务场景进行异常访问行为规则建模[7]，例如高频下载敏感数据库，敏感信息频繁外发等；也可基于日常对个人敏感信息访问行为的海量数据，建立内部用户日常行为基线，与基线比对判定是否行为异常；还可结合登录终端、地理位置信息等多种因素判定是否为异常访问。

3.1.3 零信任架构设计 在具备身份鉴别和多种访问控制措施后，可进一步构建零信任体系——通过识别用户、终端和行为，结合动态风险评估，决定是否授权访问个人敏感信息；根据访问者访问场景、权限、动作行为（如小批量高频下载个人敏感信息），结合终端环境（如是否安装杀毒软件，是否更新漏洞库与病毒库等）以及网络访问环境（是否 VPN 远程介入，是否连接无加密措施的 Wi－Fi 等）制定访问控制模型。

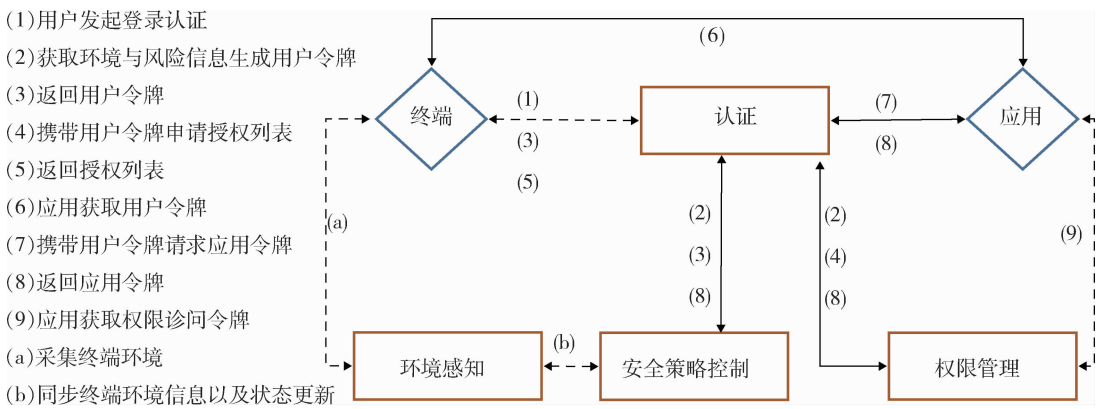

图1 零信任访问流程

基于身份的访问控制措施，优势在于可以认证个人身份并做到细粒度的访问控制，但是要对账号进行全生命周期的管理，账号与权限的映射工作较复杂。身份管控结合访问控制措施，效果更好。零信任架构要建设身份－风险识别－动态访问控制体系，其核心是身份识别，其难点是风险模型[8]，医疗卫生机构必须熟悉自身的业务和使用场景，提前将基础的安全检测、防护、身份识别措施等建设到位。

3.2 数据加密与脱敏相结合

3.2.1 数据传输加密 患者访问医疗机构的页面、下载诊疗信息、远程诊疗、线上教学等常见场景下，数据在传输过程中应当采用加密的技术手段，以防止明文传输导致的个人敏感信息泄密。数据加密传输一般在 HTTP 访问、FTP 文件访问、远程访问等场景下采用 SSL 套接字加密方式[9]进行，在不同分支机构网络间通常采用 IPSECVPN 加密方式访问。此外，如果涉及跨国机构间的科研数据共享与传输，则要在获得审批后，对出境的数据进行传输加密。

3.2.2 数据存储加密 （1）散列值加密。根据键值（key）计算哈希值，形成所需查询信息的映射；建立散列值表之后，即可以小表（散列表）驱

动大表（个人信息数据表）的方式进行查询，兼顾个人信息保护与数据库访问性能。散列值加密映射关系，见图2。散列值加密通常用于信息录入与映射，以及个人基础信息查询的场景。（2）数据库列加密。将部分个人信息字段遮蔽，以保护敏感信息。为了兼顾敏感信息保护与身份确认，列加密时对数据库结构进行优化，例如姓与名拆分形成不同的键值；又如将身份证号拆分为前6 位 - 中8 位 - 后4 位3 列。典型列加密应用场景如候诊排队叫号，遮蔽候诊患者个人敏感信息，不会造成泄密，同时又可以根据未遮蔽的部分信息进行身份确认。由于数据库管理系统无法识别库中密文，只有将整个数据库解密后才能完全查询。该加密方式影响数据库的可用性，且对每个加密列生成的密钥有较高管理要求[10]，性能与效率方面不及散列值映射方式，见表1（表内身份证信息非真实数据，仅作为示例展示）。（3）文件存储加密。多用于临床分析报告、医药实验等典型场景，考虑到存储与访问性能的均衡，一般采用对称加密与非对称加密结合的方式，密钥采用非对称算法加密，存储内容采用对称加密[11]。（4）终端加密。采用 Windows 系统的 Bitlocker 等加密手段，对终端电脑进行磁盘级别加密，保存在电脑终端的包含个人敏感信息的文件，只有通过加密终端才可访问，导出后无法查看。

个人基础信息表

序号	姓名	身份证号
1	赵一	110101197001101019
2	钱二	220221198011207 89x
3	孙三	320103199009093660
4	李四	440102200010103220

个人基础信息散列值计算后的映射表

散列值映射表

序号	个人基础信息散列值
1	Hash01
2	Hash02
3	Hash03
4	Hash04

图2 散列值加密与表空间字段映射

表1 个人信息字段列加密示例

序号	姓名	性别	身份证前6 位	身份证中8 位	身份证后4 位
1	赵一	男	110101	********	1011
2	钱二	男	220202	********	2022
3	孙三	女	330303	********	3033
4	李四	女	440404	********	4044

3.2.3 *动/静态脱敏* （1）动态脱敏。科研、教学场景中，部分授权用户访问个人敏感信息时，可以利用特殊字符遮蔽部分个人敏感信息内容，以防止泄密。动态脱敏后的效果类似前文表1。（2）静态脱敏。尤其适合开发测试场景，通过构造数据、代码转换或采用中间库/表的数据转换技术等方式进行脱敏后，再供第三方进行开发测试。静态脱敏的原理，见表2、表3，静态脱敏后个人敏感信息已经从真实信息转化为虚拟信息，实现了匿名化效果（身份证信息非真实数据，仅作为示例展示）。此时开发测试人员接触到看似数据结构一致的个人数据，但该信息已经匿名化，即便泄漏，也只是泄漏了伪造的个人信息，不会对个人真实信息造成影响与侵害。

表2 原始个人信息

字段名	数据类型	示例
User. PhoneNumber	Number	1330000000
User. Name	Varchar（20）	张三
User. ID	Varchar（20）	110000000000000001
AA01. CardNumber	Char	6225 - 0000 - 0000 - 000
AA02. Address	Text	北京市海淀区西北旺21 号

表3 静态脱敏后的个人信息

字段名	数据类型	示例
User. PhoneNumber	Number	1530000000
User. Name	Varchar（20）	李四
User. ID	Varchar（20）	220000000000000001
AA01. CardNumber	Char	6226 - 0001 - 0001 - 001
AA02. Address	Text	北京市朝阳区大屯路东15 号

此外，还可以基于代码或者使用专用软件实现对医学数字成像和通信（digital imaging and communications in medicine，DICOM）文件中的患者信息，如姓名、性别、年龄、医院信息等的匿名化处理。影像文件中个人敏感信息匿名化常用于内外部的科研与教学场景。

3.2.4 K 匿名技术 K 匿名不仅可以最大限度地保护公布数据的统计特征，而且不会泄漏公众个人敏感信息。其原理是每条记录在发布前，都要至少与 K－1 条记录无法区分开来[12]，见表 4，表中包括年龄段、性别、症状、地区 4 个特征，此时 K－1＝3。具有相同准标识符的记录构成一个等价类。准标识符指介于标识符与非敏感属性之间的一些属性，将这些属性与其他数据表结合，能够识别出患者的具体信息。如此即使攻击者具备其他背景知识，也无法推理出群体样本中个人患者的信息。

表 4 K 匿名示意

地区	年龄段（岁）	性别	疾病	邮编
四川	20～40	女	肿瘤	61 ****
四川	20～40	女	高血压	61 ****
四川	20～40	男	糖尿病	64 ****

综合上述技术，小规模访问场景可以考虑使用数据库自带的列加密方式保护个人信息。大规模访问场景可以通过散列值或数据库动态脱敏方式保证安全。用户注册与使用场景下，可以通过散列值方式，保证查询效率与安全保护的平衡。内部使用与交互、设备采集保存的场景下，可以通过列加密，数据静态脱敏，动态脱敏或 K 匿名技术保障数据安全。信息披露场景下，可以将散列值方式以及列加密/脱敏方式结合展示，保障数据安全。

3.3 水印与审计

通过设置屏幕水印为明文关联或暗码（如点阵或通过转换部分字体）的方式，可以对拍照/截屏等泄密方式进行有效震慑。且在数据泄漏后，可以利用水印对源头进行审计与追溯。

此外还有多重审计方式。一是运维审计。对存有个人敏感信息的系统运维时，应当进行运维审计，将运维人员的操作进行命令行归档/录屏，对高危命令（如删库、删表、格式化）进行告警与阻断。二是权限审计。针对账号权限，尤其是院内人员的访问权限，应当进行权限审计，细粒度地根据岗位、对应业务系统访问权限的映射关系进行评估，防止权限过大导致的汇聚推理泄密或违规访问。三是行为审计。对个人敏感信息访问行为，可以采集 Syslog、运维操作命令行归档/录屏，或通过 ODBC、FTP、API 与其他定制接口进行数据采集，结合机器学习与自定义规则进行数据访问与泄漏场景建模，识别个人敏感信息流转行为、异常访问，并对个人敏感信息来源进行追溯审计（结合水印技术、时间戳技术），定位源头。对内结合日志、运维命令行归档/录屏、水印技术等多种方式，对个人敏感信息的内部访问、流转进行电子证据存取。如发生大规模泄密事件，可利用多种证据进行多方审计。四是日志审计。Syslog 日志会将正常操作、入侵、异常告警等都记录下来，针对日志的审计往往可发现问题。通常要有日志管理平台进行相关日志的集中收集，并且满足《中华人民共和国网络安全法》日志留存 180 天的合规要求。此外，还要结合威胁情报、行为特征库等进行比对，以期通过日志分析发现入侵窃密、恶意删除、违规操作等行为。

综合上述技术，水印技术通常用于追溯桌面拍照或截图方式的泄密，而基于流量、日志、操作等方式的审计，除了法律合规要求之外，更多的意义在于最后一重底线的保障。审计往往需要耗费大量人力，未来可以结合医疗卫生机构自身的业务流程、业务场景特性以及安全基线等，形成半自动或自动化的方式，以提升效率。

4 结语

从医疗卫生机构个人敏感信息使用的典型场景出发，针对用户注册、内部利用与交互、医疗器械/可穿戴设备采集和信息披露等使用场景，分析个人敏感信息泄漏风险，结合身份鉴别、访问控制、数据加密、数据脱敏和检测审计等多种技术手段，探究不同技术手段的适用场景以及优缺点。这些技术手段可以辅助医疗卫生机构强化个人敏感信息使用安全，在医疗卫生行业中具有一定的示范性。只有规划得力、技

术与管理措施到位、人员能力匹配，适当关注和利用新技术，持续进行优化改进，医疗卫生机构个人信息保护能力才能保持在较高水平。

利益声明：所有作者均声明不存在利益冲突。

参考文献

1 中华人民共和国网络安全法［EB/OL］.［2022－11－07］. http：//www. npc. gov. cn/npc/c30834/201611/270b43e8b35e4f7ea98502b6f0e26f8a. shtml.

2 中华人民共和国民法典［EB/OL］.［2022－11－07］. https：//www. gov. cn/xinwen/2020－06/01/content_5516649. html.

3 中华人民共和国数据安全法［EB/OL］.［2022－11－07］. http：//www. npc. gov. cn/npc/c30834/202106/7c9af12f51334a73b56d7938f99a788a. shtml.

4 中华人民共和国个人信息保护法［EB/OL］.［2022－11－07］. http：//www. npc. gov. cn/npc/c30834/202108/a8c4e3672c74491a80b53a172bb753fe. shtml.

5 国家互联网信息办公室. 数据出境安全评估办法［EB/OL］.［2022－11－07］. http：//www. cac. gov. cn/2022－07/07/c_1658811536594644. htm.

6 中国信息通信研究院，深信服科技股份有限公司. 医疗物联网安全研究报告（2021 年）［R/OL］.［2022－11－07］. http：//www. caict. ac. cn/kxyj/qwfb/ztbg/20217/t20210728_381241. htm.

7 刘进，李江波，叶兵. 对于 UEBA 数据安全内控风险管理的研究［J］. 网络空间安全，2021，12（Z3）：43－48，55.

8 王斯梁，冯暄，蔡友保，等. 零信任安全模型解析及应用研究［J］. 信息安全研究，2020，6（11）：966－971.

9 孙晓叶，陈新房，杜春雷. SSL 协议的应用与安全性分析［J］. 福建电脑，2014，30（7）：59－61，84.

10 孟凯. 数据库加密系统的设计与实现［D］. 太原：太原理工大学，2012.

11 康海燕，邓婕. 面向医疗数据安全存储的增强混合加密方法［J］. 北京理工大学学报，2021，41（10）：1058－1068.

12 杨柳，李云. 混合式的 K－匿名特征选择算法［J］. 计算机应用，2021，41（12）：3521－3526.

（上接第 63 页）

15 PIERI M，CIALONI D，MARRONI A. Continuous real－time monitoring and recording of glycemia during scuba diving：pilot study［J］. Undersea & hyperbaric medicine，2016，43（3）：265－272.

16 张雅茜，张杏花，范柳媚. 白云区成人糖尿病知识知晓率及其影响因素分析［J］. 慢性病学杂志，2022，23（2）：188－190.

17 张金玲，许慧琳，张芬，等. 社区高血压与糖尿病患者核心知识知晓情况及影响因素调查［J］. 慢性病学杂志，2020，21（10）：1473－1475，1479.

18 毕浩然，刘军军，曾智. 苏北地区慢性病病人自我管理水平及其影响因素研究［J］. 护理研究，2021，35（10）：1858－1861.

19 HÖCHSMANN C，INFANGER D，KLENK C，et al. Effectiveness of a behavior change technique－based smartphone game to improve intrinsic motivation and physical activity adherence in patients with type 2 diabetes：randomized controlled trial［J］. Journal of medical internet research serious games，2019，7（1）：e11444.

20 KIRWAN M，VANDELANOTTE C，FENNING A，et al. Diabetes self－management smartphone application for adults with type 1 diabetes：randomized controlled trial［J］. Journal of medical internet research，2013，15（11）：e235.

21 李昂，郭晓蕙，张俊清. 互联网＋糖尿病共同照护模式的线上自我管理支持对规律复诊 2 型糖尿病患者代谢指标影响的研究［J］. 中国糖尿病杂志，2021，29（4）：275－278.

22 袁东登，乐嘉宜，郁敏杰，等. “微信公众号”在社区糖尿病管理中的应用效果研究［J］. 中国全科医学，2018，21（19）：2353－2356.

23 YANG Y，PENG N，CHEN G，et al. Interaction between smoking and diabetes in relation to subsequent risk of cardiovascular events［J］. Cardiovascular diabetology，2022，21（1）：14.

24 何巧，刘宇，赵芳，等. 新诊断 2 型糖尿病青年患者自我管理现状及影响因素研究［J］. 护理学杂志，2018，33（11）：26－30.

25 何雅薇，葛华英，李修英，等. 中青年 2 型糖尿病患者自我管理行为的影响因素研究［J］. 预防医学，2022，34（3）：258－262.