第 3 章

健康医疗大数据应用的安全

3.1 引言

健康医疗大数据涉及个人隐私、公共利益、国家安全，在健康医疗大数据创新应用过程中，安全是前提，也是创新应用业务发展的保障。要做好健康医疗大数据创新应用体系建设，需要关注安全层面的以下几方面内容。

1. 遵从健康医疗大数据伦理要求

健康医疗大数据创新应用在业务开展过程中需要兼顾平衡各方利益，尤其是健康医疗数据涉及高度敏感的个人隐私，需要关注伦理问题并遵从应用审查要求，这也是健康医疗大数据区别于其他领域大数据的鲜明特点。

2. 符合健康医疗个人信息保护要求

健康医疗大数据是由大量个人健康医疗信息汇聚而成的，既具有个人属性，又具有公共属性，同时也涉及国家安全，其中的个人信息保护问题非常突出，需要依法依规保护。

3. 借鉴健康医疗数据安全指南要求

健康医疗大数据创新应用的关键在于数据的流转和开放，需要借鉴有关国家行业标准，全面分析、设计各个环节和场景的安全措施要求，确保数据被安全地应用。

4. 开展健康医疗大数据网络安全保障体系建设评估

为保障健康医疗大数据创新应用的安全，需要建立健康医疗大数据应用的网络安全保障体系，并适时开展以等级保护和风险评估为重点的网络安全评估，确保网络安全保障体系的有效性。

本章主要围绕上述内容展开讲述，包括健康医疗大数据伦理、健康医疗个人信息保护、健康医疗数据安全指南、健康医疗大数据网络安全保障体系建设评估。讨论健康医

疗大数据应用的安全，必须理解健康医疗大数据伦理、健康医疗个人信息保护、健康医疗数据安全指南、健康医疗大数据网络安全保障体系建设评估等有关的概念内涵。

什么是健康医疗大数据伦理？

医学伦理是医学的一个重要组成部分。医学伦理作为一种公共意识，与现代系统医学是相伴而生、相辅相成的。健康医疗大数据的应用使得医学伦理不仅远远超越了传统伦理学的范围，更进一步扩展了医学伦理的内涵和外延，带来了数据主义思潮与医学伦理问题的挑战。

健康医疗个人信息保护如何进行？

《中华人民共和国个人信息保护法》中，把个人信息定义为以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息，不包括匿名化处理后的信息。而健康医疗信息被定义为敏感个人信息，需要采取更严格的限制，健康医疗个人信息作为敏感个人信息的聚集体，不可避免地需要更严格、更合理的合规监管政策和技术防护手段。

何为健康医疗数据安全标准？

这里的健康医疗数据安全标准主要指的是国家网络安全标准 GB/T 39725—2020《信息安全技术　健康医疗数据安全指南》。通过健康医疗数据安全标准的应用，能够确保健康医疗数据的保密性、完整性和可用性以及数据使用和披露过程的合法性和合规性，为健康医疗大数据创新应用保驾护航。

如何建设健康医疗大数据网络安全保障体系建设评估？

健康医疗大数据的创新应用，大大增加了行业业务系统的互联网暴露面，需要综合网络安全、数据安全、个人信息安全等要求建立整体的网络安全保障体系，构建体系化的安全合规评估能力，提升健康医疗大数据应用的安全防护水平，降低安全事件发生的可能性。

3.2　健康医疗大数据伦理

医学伦理是在医学基础上融合伦理学衍生出来的，研究对象是医学道德现象，即

医学领域中的道德现象。医学伦理通过对医德、医学科学道德、卫生管理道德、患者道德的研究，确定医学道德的目标，构建医学伦理学的基础理论，分析医学伦理的历史和现实，实现医学道德。大数据技术在健康医疗领域的应用不仅是一种技术变革，更催生了基于“全面数据化”的数据主义新的观念和意识形态，健康医疗有关的大数据收集和应用势必对现行的医学伦理价值产生冲击和挑战，因此，开展健康医疗大数据伦理审查显得尤为重要。

3.2.1 健康医疗大数据应用伦理问题和挑战

大数据技术在健康医疗领域的应用为卫生健康行业带来了新的伦理问题和挑战。我们在健康医疗大数据应用中必须正确认识和应对这些问题和挑战，以便更好地趋利避害。

首先，在健康医疗大数据创新应用中，必须正确认识大数据技术带来的伦理问题，主要表现在以下 3 个方面：

（1）隐私泄露问题：数字时代，人们随时随地都可能穿戴智能设备、使用智能平台，这使得个人的身份、位置、购买行为、情感、社交关系等隐私信息，都可能被真实记录、永久保存和可视呈现，而大数据环境往往是开放性环境，一旦受到攻击或被违规使用，个人隐私保护将无从谈起。

（2）网络安全问题：大数据环境下被收集利用的个人信息很容易被滥用，其所有权、知情权、同意权、保存权、修改权、使用权、开放权、删除权以及隐私权等在多数情况下都很难得到保障，大数据平台自身的安全漏洞，以及安全事件频发的数据泄露及高科技犯罪，使得网络安全必然衍生伦理问题。

（3）数据鸿沟问题：大数据技术创新应用使得一部分人能够较好地利用大数据资源，获取市场化数字红利，而另一部分人则无法利用大数据资源，这就导致数据权益分配和流通秩序的不公平，加剧社会经济差异和矛盾。

其次，在健康医疗大数据创新应用中，必须正确认识大数据带来的数据主义伦理挑战。大数据技术在健康医疗领域的运用，要求实现数据的开放和共享，给健康医疗大数据的创新应用带来了新知识，创造了新价值，提升了新能力，使得健康医疗大数据应用产生了革命性改变，这种技术变革推崇“数据和算法至上”，形成了信息自由的数据主义，这与世界主要的人本主义价值观产生了矛盾和冲突，造成个人隐私保护的隐患，容易引起数据滥用或垄断。在涉及人类生老病死的健康医疗领域，伦理挑战显得更加严重。基于数据主义考虑，个人隐私保护权利应该让位于社会整体的利益，保障健康医疗数据高效流通，而数据的算法可以不必解释其中的因果关系，只要保证

高效应用即可。

显然，健康医疗大数据创新应用带来的这种伦理问题和价值挑战不会被现实社会全盘接受，这显然会破坏人的权威和意义来源，不符合医学伦理的要求。不仅如此，这种挑战如果缺乏行业主管部门的强力制约，就会加剧未经同意的大数据收集和使用，出现更多不公平、不透明的自动化收集个人信息应用场景，以及“黑箱”算法在诊疗活动中的任性使用等现象。因此，非常有必要强化行业主管部门对健康医疗大数据应用伦理的监管，加强健康医疗大数据应用的伦理审查，弱化数据主义在健康医疗大数据应用中的滋生蔓延，守护医学伦理秩序。

3.2.2 健康医疗大数据应用伦理审查

《赫尔辛基宣言》是国际上认可接受国家最多的医学伦理审查文献。国家卫生健康委员会印发《涉及人的生物医学研究伦理审查办法》（简称《伦理审查办法》），充分吸纳了《赫尔辛基宣言》有关伦理审查研究的要求。国家食品药品监督管理局印发《药物临床试验伦理审查工作指导原则》（简称《伦理审查工作原则》），明确引用了《赫尔辛基宣言》。

《伦理审查办法》详细描述了生物医学研究伦理审查范围，规定了生物医学伦理委员会的设立要求、人员构成、审查职责、审查程序以及监督管理等一系列内容。《伦理审查办法》要求在研究中充分尊重受试者的自主决定权，同时遵守风险可控、公平合理、保护隐私、依法赔偿、特殊保护等原则，重点对知情同意有关事项做出了详细规定，要求知情同意书必须包含必要完整的信息和内容，并按照规范的流程获得受试者的知情同意。《伦理审查办法》同时也对研究者何时需要再次获取知情同意书，以及经审查批准后可免除签署知情同意书等情形提出了明确要求。

《伦理审查工作原则》从伦理委员会对药物临床试验的科学、合理审查出发，首先明确了伦理审查的组织管理与职责要求，其次规范了伦理审查的申请受理、开展审查、决定送达、跟踪审查等工作流程，最后规定了伦理审查的文件管理要求。《伦理审查工作原则》通过提出药物临床试验伦理委员会审查的重点，旨在提高伦理委员会的伦理审查能力，规范药物临床试验伦理审查工作。

基于以上医学伦理审查原则，健康医疗大数据创新应用在运用大数据技术时，需要按照医学伦理原则积极开展大数据应用伦理审查工作：①需要建立医学伦理委员会，委员应包含不同性别的成员，并由生物医学、伦理学、社会学、法学等领域专家以及外部机构社会人士组成，人数不少于 7 人；②需要建立科学合理的医学伦理审查工作制度规程，保证伦理审查过程的合法、公正、客观、独立；③需要明确医学伦理

委员会的职责和义务，保证受试者的尊严、权益及安全，促进生物医学伦理研究的规范开展。伦理委员会应采取相关利益冲突防范机制，保证伦理审查工作的独立性。同时，相关行业主管部门也需要基于健康医疗大数据的技术特征建立健全新的伦理审查规范，促进健康医疗大数据技术的良性应用和发展。

（3.2 节作者：金涛）

3.3 健康医疗个人信息保护

个人信息是以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息。数字经济时代，随着大数据、人工智能等新技术在健康医疗领域的应用，大数据系统中的数据安全保护难度加大，个人信息泄露风险加剧，个人信息所有权难以保障，用户隐私泄露问题变得日益严重。一方面，数据的过度收集和集中处理导致了个人信息的滥用、泄露、非法交易等；另一方面，多源数据关联分析正在严重威胁用户隐私。在健康医疗大数据应用中是要把大数据等新技术当作洪水猛兽而拒之门外，还是在现行法律框架下，动态平衡个人信息保护与医疗数据应用之间的冲突，值得我们思考。

3.3.1 国内外健康医疗个人信息保护法律法规

健康医疗个人信息是指健康医疗数据中的个人信息，属于敏感个人信息，健康医疗个人信息保护是一个全球性的问题，许多国际公约和国家宪法明确将个人信息保护作为核心原则或目标。

欧盟、美国、澳大利亚、日本等世界发达团体和国家相继立法规范健康医疗个人信息保护。2016 年，欧盟印发《通用数据保护条例》（General Data Protection Regulation，GDPR），将健康数据被归类到“特殊类别”，要求对个人健康信息进行严格保护，无论健康数据是否涉及个人隐私，非法律明确承认的理由，各组织均不得处理个人数据，即使是公开可获取的数据，只要涉及个人信息均在该条例的保护范围内。美国针对个人健康信息保护发布的最主要法律是《健康保险携带和责任法案》（Health insurance portability and accountability Act，简称 HIPAA 法案），与欧洲的情况相反，HIPAA 法案主要目的在于保护个人医疗数据隐私的同时提高相关实体医疗的治理水平和效率，对于在治疗过程中产生的个人健康信息的保护权、知悉权、信息共享等场景都做了详细的安全防护规定，来确保患者的个人隐私。澳大利亚非常重视电

子健康档案隐私保护，在 2016 年成立了澳大利亚电子健康署（Australian Commission for Electronic Health），并颁布了全国性的电子健康档案发展战略，详细分析了电子健康档案的隐私泄露风险，构建了电子健康档案数据全生命周期的隐私保护体系。日本在 2003 年颁布《个人信息保护法》，规定在有效利用个人信息的情况下对个人信息进行保护，并在 2014 年进行战略补充，鼓励在医疗大数据平台下灵活有效地利用医疗数据，提升医疗服务效率，提倡以数据分析利用达到降低医疗费用的目的。

“十三五”以来，我国出台了一系列法规标准和政策文件来规范健康医疗个人信息保护，主要包括以下 3 个方面：

（1）法律法规方面：出台了包括《中华人民共和国网络安全法》《中华人民共和国民法典》《中华人民共和国数据安全法》《中华人民共和国个人信息保护法》等在内的一系列政策法规，明确了个人信息保护的通用基础合规建设要求。

（2）安全标准方面：出台了包括 GB/T 35273—2020《信息安全技术　个人信息安全规范》、GB/T 39725—2020《信息安全技术　健康医疗数据安全指南》、GBT 37964—2019《信息安全技术　个人信息去标识化指南》等在内的一系列的国家标准规范，明确了健康医疗个人信息在收集、存储、使用、共享、披露等环节的行为和防范措施。

（3）行业政策方面：国家和行业主管部门先后发布《关于印发国家健康医疗大数据标准、安全和服务管理办法（试行）的通知》《关于印发加强网络安全和数据保护工作指导意见的通知》等政策文件，加强行业数据安全和隐私保护能力。

通过对国内外健康医疗个人信息保护法律法规的研究可以发现，健康医疗个人信息需要按照合规遵从原则进行保护，先将健康医疗领域适用的外部法规标准进行分解重组，形成行业内部的“合规基准”，基于“合规基准”开展安全治理，形成行业性的政策标准和技术规范。随着《中华人民共和国个人信息保护法》正式实施，个人信息保护立法体系已经建立，但行业“内部基准”还未形成，现有的政策文件主要关注数据安全，如何充分研究健康医疗个人信息的保护原则和特殊性，将个人信息保护与健康医疗领域紧密结合，形成健康医疗个人信息的政策标准和技术规范，还任重道远。

3.3.2　个人信息保护基本原则

个人信息承载着多重利益，不仅有自然人的人格尊严、隐私权及个人信息权益等民事权益，也有企业、国家机关等主体合理利用个人信息的利益，还涉及言论自由、公共安全、国家安全等。因此，需要依法依规明确个人信息保护的基本原则，在维护好个人信息权益的同时，促进信息数据合理有效利用。

目前，我国已出台的多部法律法规对个人信息保护应遵循的原则均有明确规定。① 2012 年颁布的《全国人民代表大会常务委员会关于加强网络信息保护的决定》第 2 条规定：“网络服务提供者和其他企业事业单位在业务活动中收集、使用公民个人电子信息，应当遵循合法、正当、必要的原则”；② 2017 年颁布的《中华人民共和国网络安全法》第 41 条规定：“网络运营者收集、使用个人信息，应当遵循合法、正当、必要的原则”；③ 2018 年正式实施的《信息安全技术　个人信息安全规范》（GB/T 35273—2017）明确个人信息控制者应遵循权责一致、目的明确、选择同意、最小必要、公开透明、确保安全、主体参与的原则；④ 2020 年颁布的《中华人民共和国民法典》第 1035 条明确要求：“处理个人信息的，应当遵循合法、正当、必要原则，不得过度处理”；⑤ 2021 年颁布的《中华人民共和国个人信息保护法》第 5 ~ 第 9 条分别规定了个人信息处理的合法、正当、必要与诚信原则、目的原则、公开透明原则、质量原则和责任原则。

《中华人民共和国个人信息保护法》颁布的 5 项原则基本集成了有关法规标准中对个人信息处理活动的原则要求，因此，在健康医疗大数据应用场景下，健康医疗个人信息控制者开展个人信息处理活动，应在法律法规框架引导下遵循合法、正当、必要与诚信原则、目的原则、公开透明原则、质量原则和责任原则 5 项保护原则。

3.3.3 健康医疗个人信息合规保护实现路径

《中华人民共和国个人信息保护法》明确规定，医疗健康数据属于敏感个人信息。健康医疗个人信息作为医疗健康数据的最基础组成部分，必须严格按照敏感个人信息的处理规则建立合规保护体系，进行合规遵从保护。健康医疗大数据组织可以将数据安全能力作为基础，以数据安全和个人信息保护合规要求为抓手，做好健康医疗个人信息保护，具体的合规保护实现路径如图 3-1 所示。

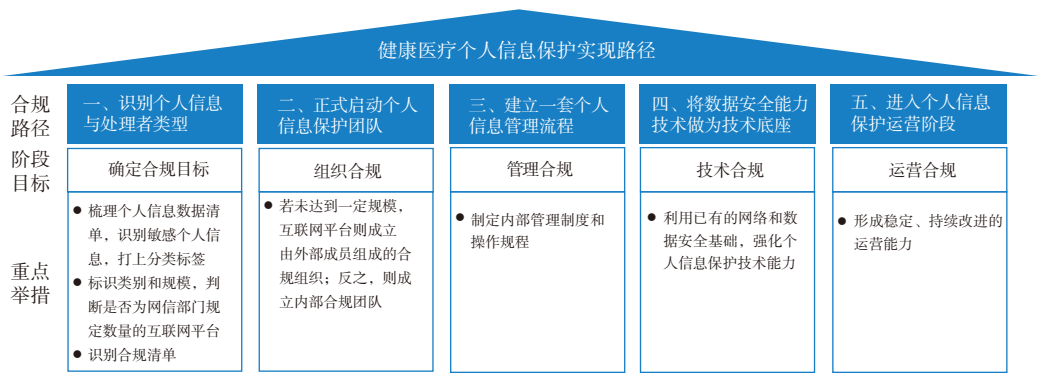

图 3-1　个人信息合规保护实现路径

（1）健康医疗组织需要建立个人信息数据清单，通过分类分级，确定需采取的安全保护措施。同时确认个人信息的数据量，以判断是否达到了网信部门规定的个人信息数据量规模，是否提供重要互联网平台服务、用户数量巨大、业务类型复杂的个人信息处理者，明确需满足的合规要求及可能存在的问题，确定合规目标。

（2）不同类型和规模的个人信息处理者有不同的组织要求，健康医疗组织应依据合规目标，明确个人信息合规保护责任部门与人员类型，组建个人信息保护团队。

（3）健康医疗组织需尽快开展制定适合于大数据业务特点的管理要求和业务场景，通过细化操作规程，管控个人信息安全风险，确保对个人信息主体权益的影响合规。

（4）健康医疗组织需要“技管并重”，以数据安全能力为技术底座，充分利用网络安全及数据安全防护产品，通过技术防范个人信息在个人信息处理活动中的风险。

（5）在保护团队、管理措施和技术措施健全后，个人信息保护进入体系化运营阶段，形成稳定的、可持续的健康医疗个人信息合规保护体系，确保个人信息安全。

3.3.4 健康医疗个人信息保护特殊性

与一般个人信息相比，健康医疗个人信息具有特殊性。

1. 健康医疗个人信息具有高度隐私敏感性

医疗数据记录和描述患者个人健康和生理状况，患者个人有权维护自己隐私权不受侵犯。患者隐私权不仅包括患者在诊疗过程中向医生和医院披露的个人基本信息（姓名、年龄、住址、电话）、病因信息（个人或家族病史等与疾病有关的情况）、生理信息等，还包括病历、标本等患者记录。涉及患者的个人信息一旦遭到泄露并被不法使用，不仅可能给患者造成经济损失，还可能给患者造成名誉损害和精神损害。在新冠病毒感染疫情防控中也发生多起泄露患者个人隐私数据的情况，既增加了社会上对疫情无意义的恐慌，又对确诊患者的正常生活和正当权利产生了不良影响。

2. 健康医疗个人信息涉及公共福祉

随着健康医疗数据加工水平的提升，基于大数据分析能够实现疾病辅助诊断、药物研发、疾病预防等多方面应用，健康医疗数据决定着医疗领域的发展，医生及行业研究人员均需要借助医疗数据才能更好地服务患者，包括改进诊断、制定治疗方案和寻找新的治疗方法等。关于个人信息保护相关立法的目的不仅在于“保护”，而是“保护”和“利用”同步推进，卫生健康管理部门基于公共利益和合法目的，可以调取和使用健康医疗数据，充分分析挖掘数据的社会公共价值，促进建立良好的健康医疗数据治理体系，实现在健康医疗大数据环境下数据资源的社会化利用。

3. 健康医疗个人信息有关国家安全及种族利益

以电子病历为核心的医疗信息化建设显著提升了医疗机构的医疗科研水平，凸显了健康医疗数据的资源性，进一步推进了健康医疗数据对个人信息数据采集的广度和深度，大集中的医疗数据具有很强的社会性、公共性，甚至会对国家安全和种族安全造成影响。国家科技部下发《人类遗传资源管理条例实施细则》规定，不得向境外提供我国人类遗传资源。国家卫生和计划生育委员会研究制定的《人口健康信息管理办法（试行）》明确规定，不得将人口健康信息在境外的服务器中存储，不得托管、租赁在境外的服务器。

对个人信息的隐私保护是数据安全利用的前提，在医疗信息化快速发展的时代，维护国家安全和种族安全是我国每个公民应尽的义务，应不断研究探索在法律法规及监管框架下最大化发挥医疗数据资源价值。

4. 大数据背景下健康医疗个人信息保护技术

大数据背景下，健康医疗个人信息保护必须依靠规范化的技术手段，有效遏制个人信息的收集利用，合理支持个人信息的删除权限，确保个人信息安全。典型的个人信息保护技术手段是去标识化。去标识化是指个人信息经过处理，使其在不借助额外信息的情况下无法识别特定自然人的过程。健康医疗大数据应用采用去标识化方法一般需要确定具体的去标识化模型和去标识化技术。常见的去标识化模型包括 K- 匿名模型和差分隐私模型等；常见的去标识化技术包括统计技术、密码技术、抑制技术、假名化技术、泛化技术、随机化技术、数据合成技术等，如图 3-2 所示，具体模型和技术解释可参考 GB/T 37964—2019《信息安全技术　个人信息去标识化指南》进行了解。

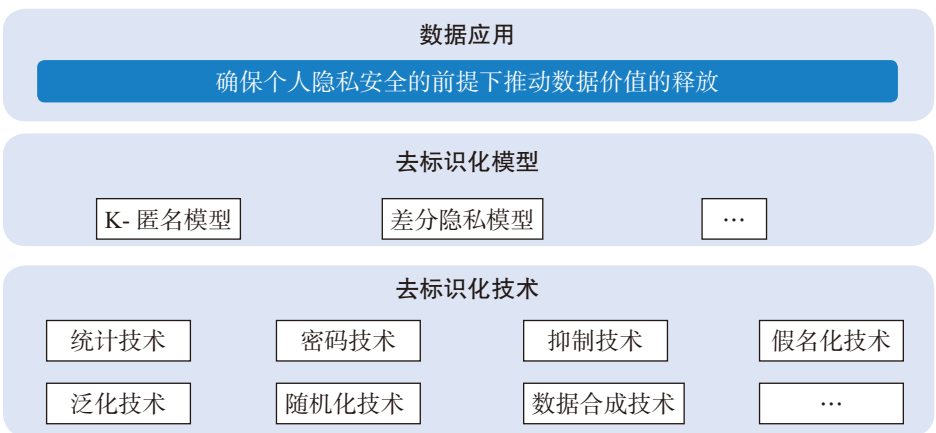

图 3-2　去标识化模型和技术

（3.3 节作者：景鸿理　张志　屈伟）

3.4 健康医疗数据安全指南

健康医疗大数据创新应用需要具有行业特点的国家安全标准的支撑，GB/T 39725—2020《信息安全技术　健康医疗数据安全指南》作为兼顾行业应用和网络安全的国家标准，为健康医疗数据提供了安全底线。通过标准实施的规范、引领和支撑作用，能够快速推进互联网、大数据、人工智能等新兴信息技术与卫生健康行业的创新融合安全发展，为健康医疗大数据创新应用的安全实践提供指导。

3.4.1 健康医疗数据安全指南标准

2021 年 7 月 1 日，国家网络安全标准 GB/T 39725—2020《信息安全技术　健康医疗数据安全指南》（以下简称《指南》）正式实施，这是我国首部针对健康医疗数据安全工作的国家标准，填补了我国健康医疗数据安全标准的空白。

《指南》是在“互联网 + 医疗健康”、智慧医院蓬勃发展的背景下发布的。健康医疗数据安全事关患者生命安全、个人信息安全、社会安全和国家安全，随着各种新业务、新应用的不断出现，健康医疗数据在全生命周期各阶段均面临着越来越多的安全挑战，安全问题频发。为更好地保护健康医疗数据安全，特制定该标准。

《指南》为推荐性国家网络安全标准，标准发布的宗旨是为健康医疗数据处理活动提供了指南，可以在保护健康医疗数据安全的前提下，规范和推动健康医疗数据的融合共享、开放应用，促进健康医疗事业发展。

《指南》发布的目的是给出健康医疗数据控制者在保护健康医疗数据时可采取的安全措施，适用于指导健康医疗数据控制者对健康医疗数据进行安全保护，也可供健康医疗、网络安全相关主管部门以及第三方评估机构等组织开展健康医疗数据的安全监督管理与评估等工作时参考。

《指南》围绕健康医疗数据业务，提出了健康医疗数据使用和披露的原则要求，解决健康医疗数据使用的安全合规边界问题；围绕数据安全措施，给出了健康医疗数据分类分级以及各级安全要点、使用场景分类以及各类场景安全要点、开放形式分类以及不同开放形式安全要点、安全管理指南（包括组织保障、PDCA、应急体系）、安全技术指南（包括通用安全指南、去标识化指南）；围绕各种常见典型场景数据，给出了安全重点措施。

《指南》并不是孤立地发挥作用，需要结合其他网络安全标准共同发挥作用，《指南》侧重于数据安全层面，更多地偏向于健康医疗数据业务层面。涉及健康医疗信息