“数据融合”的风险提示与合规要点

作者：潘永建 | 邓梓珊 | 左嘉玮

大数据时代，收集用户数据的渠道日益增多，多源数据融合已成为企业运营和发展的关键议题。通过将不同来源的数据融合在一起，企业可以更加精准地刻画用户的画像，从而获得更好的业务洞察能力。目前，数据融合已在零售业、金融行业、医疗和等多个行业进行落地实践，依据数据进行了更多场景的孵化，如识别潜在客户、定制市场营销等。

相对于传统的基于单一数据源的分析利用，数据融合能够产生更大的商业价值，但是海量数据的汇聚和分析可能导致许多新的法律风险。本文将分析数据融合所面临的特殊法律挑战以及相应的合规建议，以供企业参考。

数据融合的定义

数据融合这一概念最早出现于 20 世纪 70 年代，是指将来自多传感器或多源的信息和数据以一定的标准进行自动分析和综合，以获得更为准确的结果。大数据时代，数据融合的内涵变得更为广泛，其定义可以概括为“将多种数据源中的相关数据提取、融合在一起进行分析和统计，挖掘出有意义和有价值的信息，并依此作出更为科学的决策”。

对于企业而言，数据融合的应用场景主要可以分为两种：

- 企业内部的数据融合：例如，企业通过数据中台打通底层数据，将不同业务线之间所产生的数据汇聚在一起，实现共享并在此基础上加工、分析、复用。

如您需要了解我们的出版物，
请联系：

Publication@llinkslaw.com

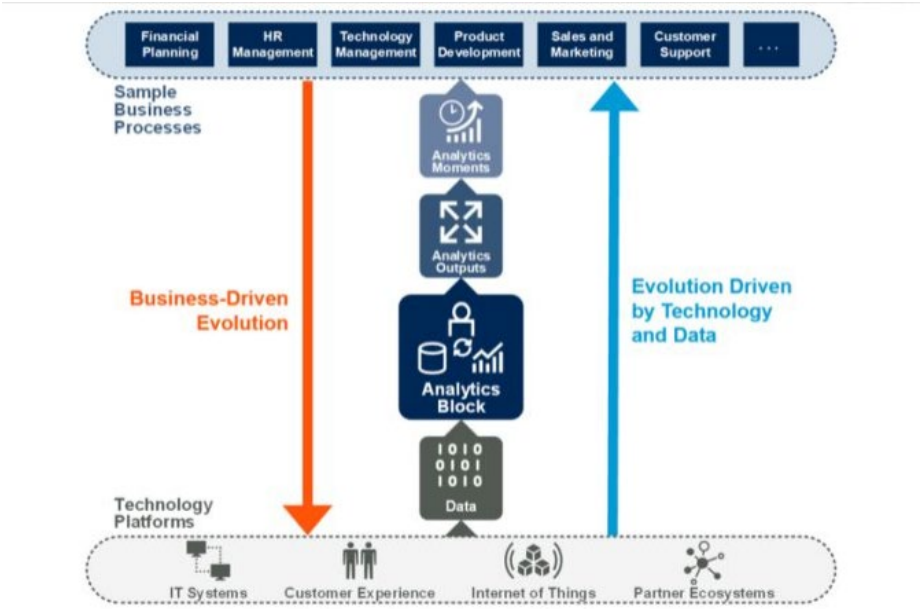

图 1-数据中台架构

来源: Gartner

- 企业与外部第三方的数据融合： 例如， 金融机构将自有数据与电商、医疗等其他行业的数据融合，实现联合建模，构建更加精准的营销和风控模型。

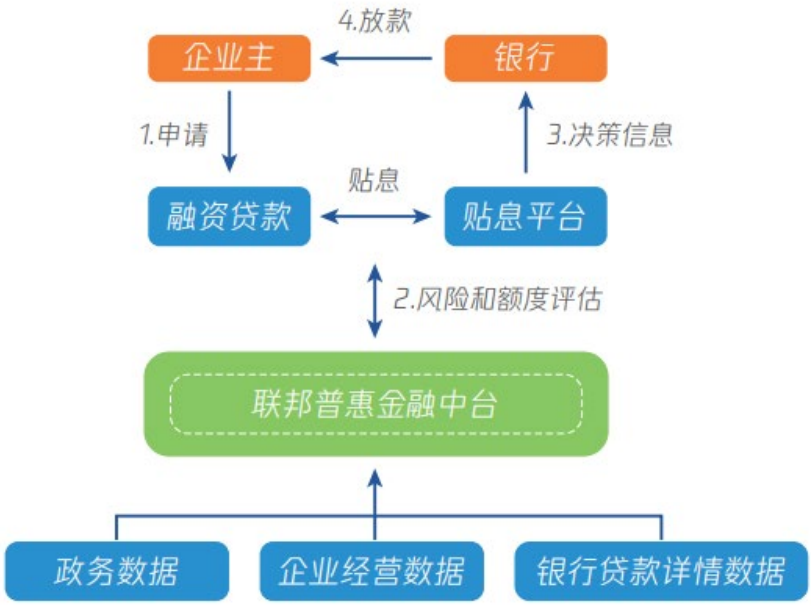

图 2-联合建模实现多方数据融合

来源：腾讯《隐私计算白皮书》

数据融合的法律风险

按照《信息安全技术个人信息安全规范》，个人信息的汇聚融合应当遵循个人信息使用的目的限制等要求。[1]个人信息的融合无疑属于一种个人信息的“处理”活动，因此在数据融合过程中，若涉及用户的个人信息，必须受到个人信息保护原则的约束。与传统的个人信息处理相比，数据融合的特殊法律问题体现在以下几个方面：

用于融合的个人信息来源合法性存疑

目前，关于判断个人信息来源的合法性存在许多争议和存疑的问题。例如，用户的授权同意是企业处理个人信息的主要法律基础，而依据我国《个人信息保护法》(以下简称“《个保法》”)的要求，同意应当由个人在充分知情的前提下自愿、明确作出。[2]因此，企业应当告知用户数据融合的目的、具体的处理方式和范围，同时该目的应当明确、合理。[3]然而，许多数据的价值和潜在用途往往在汇聚融合之后才会变得清晰，企业很难提前预测并告知用户具体特定的处理目的。在实践中，企业往往会在个人信息保护政策中使用“为了提供更好的服务，我们会与第三方共享您的个人信息”之类笼统的描述，这很难说是一个被监管部门充分接受的目的。2009 年 1 月，CNIL(法国数据保护执法机关)就以“个性化广告”之处理欠缺合法基础为由，对 Google 处以 5000 万欧元罚款。CNIL 指出，谷歌将大量数据进行组合和分析(数据融合)，这种数据处理行为会对用户权益造成显著影响，但是谷歌在隐私政策中对于数据处理目的的描述过于笼统和模糊，未以合理的“透明度”告知用户，并不能获得用户的有效同意。[4]

数据融合削弱去标识化效果

由于《个保法》将(真正的)匿名化数据排除在个人信息范畴之外，企业往往希望通过技术手段对数据进行匿名化，以豁免自身的合规义务。但遗憾的是，大数据带来的多源数据融合使得传统的去标识化技术很难达到“匿名化”的效果，难以对个人信息主体进行有效保护。实践中，许多企业认为去除个人信息中的直接标识符(姓名、身份证号等)就可以实现匿名化要求，并在个人信息保护政策中向用户承诺，他们的数据只会以未识别的匿名方式与其他人共享。但事实上，数据融合结合大数据分析、机器学习等技术，可以将本身可能不敏感的各种数据进行关联，重新识别出特定的个人，从而无形中泄露了用户的个人信息甚至是敏感信息。这导致许多企业通常采取的“去标识化”数据保护措施，并不能达到匿名化的效果以免除相关的个人信息合规义务。

融合过程的安全性面临威胁

我国《数据安全法》(以下简称“《数安法》”)《个保法》都对数据处理活动中的安全性提出了要求[5]，企业在数据融合过程中应当采取相应的技术措施和其他必要措施，保障数据安全。考虑到数据来源的多样

[1] 参见《信息安全技术个人信息安全规范》第 7.6 条
[2] 参见《个人信息保护法》第十四条
[3] 参见《个人信息保护法》第六条
[4] The CNIL's restricted committee imposes a financial penalty of 50 Million euros against GOOGLE LLC, https://www.cnil.fr/en/cnils-restricted-committee-imposes-financial-penalty-50-million-euros-against-google-llc
[5] 参见《数据安全法》第二十七条；《个人信息保护法》第五十一条

性以及收集、汇总和存储数据的巨大数量，数据融合不可避免地会带来更高的数据安全风险，包括对数据未经授权的访问和意外泄露等。对于企业来说，如何确定与数据融合目的、处理方式和范围相适应的安全措施将是一项具有挑战性的任务。此外，在多源数据融合的场景下，往往难以确定个人信息具体是在哪个环节发生了泄露，因此企业还应对合作主体的安全措施进行验证，这进一步加重了企业的注意义务。

如果在数据融合过程中(尤其是在使用云计算服务时)需要将数据传输到中华人民共和国境外，企业还应当注意我国《数安法》《个保法》对于重要数据和个人信息出境的特殊监管要求，如安全评估、认证等。

融合结果的准确性有待加强

依据《个保法》的要求，企业在处理个人信息时应当保证个人信息的质量，避免因个人信息不准确、不完整对个人权益造成不利影响。[6]由于数据融合需要从多个不同的数据源收集数据，一方面，每一来源数据的准确性和相关性有待核实；另一方面，由于融合后的数据通常具有多个维度，不同维度的误差会产生累积，这些因素都容易导致数据分析和预测的结果出现偏差。

值得注意的是，数据融合的结果往往应用于企业的自动化决策，错误的结果可能会损害用户的公平交易权等权益，从而导致企业承担赔偿责任。例如，不同的医疗机构可以通过数据融合进行联合建模，实现AI 诊断治疗。不正确的数据可能会导致错误的诊断，从而危及用户的生命健康。

合规建议

针对上述法律风险，我们建议企业在数据融合时应注意以下合规要点：

1. 明确数据的具体类型和性质。许多类型的数据受到法律和法规的特别限制，例如敏感个人信息、人口健康信息、金融数据、重要数据等。在数据融合之前，应结合融合的目的、范围、方式等因素，判断数据是否落入某一特定的数据类型，以明确所适用的法律规定，在此基础上实现数据的分级分类保护。

2. 审查数据融合的必要性，在满足业务需求的前提下最小化数据的使用。在融合时宜对数据进行标签化处理，以减少直接使用原始数据带来的个人信息风险，并对用户标签的颗粒度进行控制，尽量避免精确定位到个人。

3. 在数据融合之前，根据融合后可能产生的数据内容、处理目的、范围、对个人的影响等对可能的个人信息风险进行评估，并对处理情况进行记录。[7]一方面，风险评估可以满足法律规定的个人信息保护义务；另一方面，加强风险沟通有助于提高数据融合的透明度，增加用户对企业的信任。

[6] 参见《个人信息保护法》第八条
[7] 参见《个人信息保护法》第五十五条

4. 审查个人信息保护政策并按需进行更新，针对不同类型的数据，应逐一列出数据融合的目的、具体的处理方式和范围、合作方的身份等，不应使用概括性的语言。若融合超出了数据收集时所声明的使用范围，应重新征得用户同意；对于敏感个人信息，应取得用户的单独同意。若数据融合涉及任何自动化决策，应当进行清晰的说明，并为用户提供合适的救济方式。

5. 数据融合的多个参与方可能被认定为共同个人信息处理者，各方应通过合同等方式明确用于融合的数据内容和范围、结果用途、各自的权利和义务，以及数据保护要求等。责任条款应当涵盖数据泄露、未经授权的访问、数据传输中断、错误数据的提供等常见的安全事件。此外，企业应明确规定是否允许融合结果的再许可以及相应的条件。

6. 应采用符合国家标准的密码技术保证数据在传输过程中的机密性，并对数据的完整性、一致性进行校验。通过 API(应用程序接口)调用形式实现数据融合的，应对接口增加鉴权机制，并实时监控调用信息，以减少恶意数据获取、数据盗用等风险。

7. 接收第三方数据时，应确认数据的来源是否合法以及其使用是否存在限制。例如，应确认数据源公司是否在个人信息保护政策中明确记载了对外合作共享的情形，并且取得了用户的授权同意。此外还应审查可能涉及数据使用和披露的任何协议，以确保第三方的数据传输不以与其合同上的义务相冲突。

8. 向第三方传输数据时，应审查数据接收方使用数据的目的、使用方式是否合法、正当、必要，是否对数据进行了分类管理并采取了相应的安全措施，并确保特殊数据(个人信息、敏感数据、重要数据等)将受到更高级别的保护。

9. 宜采用隐私计算技术降低数据融合过程中的风险，防止个人信息、敏感数据和重要数据的泄露、篡改、丢失。例如，多方安全计算、联邦学习等技术有助于最大限度地减少不必要的数据使用；可信执行环境技术可以用来建立对数据源的信任，保证收集数据的准确性和可靠性。

如您希望就相关问题进一步交流，请联系：

潘永建
+86 21 3135 8701
david.pan@llinkslaw.com

如您希望就其他问题进一步交流或有其他业务咨询需求，请随时与我们联系：**master@llinkslaw.com**

上海	北京	深圳
上海市银城中路 68 号 时代金融中心 19 楼 T: +86 21 3135 8666 F: +86 21 3135 8600	北京市建国门北大街 8 号 华润大厦 4 楼 T: +86 10 8519 2266 F: +86 10 8519 2929	深圳市南山区科苑南路 2666 号 中国华润大厦 18 楼 T: +86 755 3391 7666 F: +86 755 3391 7668
香港	伦敦	
香港中环遮打道 18 号 历山大厦 32 楼 3201 室 T: +852 2592 1978 F: +852 2868 0883	1/F, 3 More London Riverside London SE1 2RE T: +44 (0)20 3283 4337 D: +44 (0)20 3283 4323	

www.llinkslaw.com

Wechat: LlinksLaw

本土化资源 国际化视野