数据保护三大实务问题与合规建议

■ 王融

当前，对于用户个人数据保护合规问题，不论是企业内部法务还是外部律师正面临越来越多的困惑。一方面，国内法律规定十分原则，法律适用产生许多模糊地带；另一方面，云计算、大数据等新兴业务也让数据保护合规问题变得更为复杂。本文就实务中最为突出的三大问题提出分析建议。

问题一：如何判断是否为个人数据?

个人数据是个人数据保护法的核心概念。判断特定数据是否属于个人数据，是决定是否履行合规义务的第一步。如为非个人数据，则没有适用数据保护法律的必要。反之相反。

现代各国个人数据保护法（包括我国在内）对于个人数据的界定，基本上采纳“识别说”。个人数据是指与个人相关的，能够识别特定自然人的信息。这其中包括直接识别和间接可识别。在实务中，对于直接识别的个人数据一般分歧较少，如个人姓名、家庭住址、电话号码、电子邮箱地址等；分歧主要在间接可识别。因为这一开放式的“界定”并没有明确间接可识别的标准。而在云计算、大数据的冲击下，间接可识别成为一个相当模糊的地带。

反映在实务中，搜索关键词、IP 地址、cookie 等数据是否属于个人数据成为争议焦点。以用户搜索关键词为例，一般认为，搜索记录是匿名的，是非个人数据。例如用户仅仅是在搜索栏中搜了“鲜花”，用该数据很难去定位个人。然而如果用户的搜索词是高度特定的词汇，或者是将该用户的搜索词复合起来，该用户可能将会被识别。2006 年，美国在线 AOL 公司公布了 2 千万个搜索记录，且公布这些数据时，AOL 公司已经将其做了匿名化处理。然而纽约时报记者很快发现这些搜索记录中至少有一部分可以很容易地被重新识别。最终 AOL 对其公布数据的行为表示道歉。同样，对于用户 cookie 的属性判断也争议连连。在 2015 年某用户诉百度利用搜索关键词投放广告侵犯隐私一案中，对于 cookie 信息的性质认定，南京基层法院和南京中级法院作出了截然相反的判断。

合规建议：

针对个人数据边界模糊问题，美国著名数据保护法律专家Paul M. Schwartz 与 Daniel J. Solove 在 2011 年提出了个人信息 2.0 及合规体系（Personally Identifiable

Information，PII2.0）。也就是说，目前各国个人数据保护法中对个人数据的法律界定可以被看作 1.0 版本，它与当下的网路环境存在着诸多不适应，因此在其基础之上，学者提出 2.0 版本。PII2.0 的思路后被 2013 年世界经济论坛继承、发展 。目前各国应对大数据发展形势而对个人数据保护立法的修订，不同程度上也吸收了这一思路。笔者认为，在我国个人数据保护法律较为原则笼统的背景下，PII2.0 概念及合规体系具有内在合理性，能够为企业合规提供参考。

依据 PII 2.0 概念及合规体系，可以将企业收集处理的数据分为三类：

第一类，非个人数据（non-PII）。即此类数据完全与个人无关，不适用个人数据保护法。如天气气象、环境监测、地理测绘、总体性的人口数据等。

第二类，已识别个人身份数据（personally identified information）。此类数据完全适用个人数据保护法。如姓名，家庭住址、电话号码等能够确定识别、关联到特定个人的数据，需符合个人数据保护法全部合规要求，包括知情同意，允许用户访问和更正，数据处理正当合法、目的限制、保障安全等。

第三类，可能识别个人身份的数据（personally identifiable information）。此类数据结合业务场景，灵活适用个人数据保护法。如业务场景中，识别风险较高，可按照第二类数据的合规性要求处理，需满足全部合规要求；如识别风险较低，则可选择部分适用。因为，在部分业务场景下，为满足知情同意、访问与更正等合规要求，如果企业需要将可能识别的数据转化为确定可识别的数据，以联系告知用户其享有这些权利，这样反而会降低用户隐私保护。因此，对于可能识别的信息，有时仅需要满足部分合规要求，如保障数据安全，数据处理规则透明原则即可。也正是如此，目前正在制定过程中的欧盟数据保护总规中规定：企业（数据控制者）不能仅仅是为了符合个人数据保护法的要求，而被迫需要收集更多的信息以识别特定的用户（数据主体）。

需要说明的是，尽管上述三类数据的合规边界是相对清晰的，但对于特定数据的判断，应当是“动态”的，而不是“静止”的。个人数据的判断应当依据特定的业务情境，而不是预先对数据的性质作出判断。对于同一数据的法律界定，在不同的业务场景下可能会有不同的结果。这其中需要考虑的业务情境因素包括：数据本身的类型、数据处理涉及的实体、服务提供商的信任水平、收集方法、设备环境、应用和使用、以及各方之间的价值交换。这区别于传统个人数据保护法对个人数据的界定方法，它不是非黑即白的简单判断标准，而是基于一套规则动态界定。

总结来看，PII2.0 概念及合规体系对于解决数据保护合规问题提供了一种思路，但它同时对企业法务提出了更高的要求。合规工作不仅需要考虑法律规则，还需要与市场、技术团队紧密配合，深入了解产品业务线的数据处理，共同为业务合规提供解决方案。

问题二：适用“明示同意”还是“默示同意”？

我国个人数据保护法律尽管规定了收集使用个人数据应当符合知情同意原则，但并未明确：究竟是适用“明示同意”，还是“默示同意”？

对于营销活动，现行法律却同时认可了“选择加入（opt - in）与“选择退出（opt - out）两种模式。《全国人大关于加强网络信息保护的决定》规定：“任何组织和个人未经电子信息接收者同意或者请求，或者电子信息接收者明确表示拒绝的，不得向其固定电话、移动电话或者个人电子邮箱发送商业性电子信息。”

此外，一些行业的数据保护法规，出于对用户重要数据的保护目的，法规要求获得用户“书面同意”。如 2013 年《征信业管理条例》第十四条第二款规定：“征信机构不得采集个人的收入、存款、有价证券、商业保险、不动产的信息和纳税数额

信息。但是，征信机构明确告知信息主体提供该信息可能产生的不利后果，并取得其书面同意的除外。”这一规定实际也并未彻底解决明示还是默示同意的问题。因为结合我国《合同法》规定，书面形式包括了数据电文形式。也就是说，用户不论是在合同协议中签字，还是在网络电子方式“打勾”，都可被视为书面形式。而在我国商业实践中，以制定隐私条款、业务协议等方式，超出业务服务目的和范围，一揽子取得对用户个人数据收集利用授权的，已经成为普遍现象。“明示”与“默示”同意原则的区分正在模糊化。

合规建议：

从严格合规以及消减数据安全风险，实现业务健康良性发展的目的出发，建议企业对同意授权模式做出区分：

企业收集、处理的用户数据可区分为用户一般数据和用户敏感个人数据。判断的标准是后者的泄露将会给用户财产与精神利益带来更为严重的后果，而参照国际惯例，敏感个人数据一般包括反映个人财产财务、健康医疗、政治及宗教信仰、性取向、位置状况的数据以及未成年人个人数据。企业也可根据自身业务场景、将敏感程度较高的数据列入这一类型。

对于个人一般数据，建议适用“默示同意”原则。即可在提供商品或服务过程中，可推定用户同意收集其为实现业务目的的相关数据，但同时为用户提供退出通道。

对于个人敏感数据，建议适用“明示同意”原则，且区别于目前实践中一揽子授权协议方式，对于个人敏感数据，应当真正做到“明示同意”。也就是说，应当经专门特定的申请，向用户揭示提供该信息的风险，让用户可以做到单独、单次授权 。

问题三：如何对匿名化数据进行合规利用?

大数据时代，数据的采集、分析与利用逐渐构成企业运营的核心。在这一过程中，数据匿名化成为商业利用的关键。根据我国现有法律规定，我国法律严格禁止公民个人数据的出售行为。最新颁布的刑法修正案(九)也继续重申强化了出售、非法获取个人信息罪。

在此背景下，匿名化数据是否还适用于个人数据保护法？匿名化数据利用的合规要点是什么？成为商业实务中最为关切的合规问题。

合规建议：

1. 对于用户个人数据的商业利用推荐采取匿名化

企业内部的数据分析、画像活动，以及其他商业利用，在可行的情况下，推荐采用匿名化方式。数据匿名化后一定程度上可豁免个人数据保护合规义务，也可整体上降低信息安全风险。

2. 商业利用匿名化数据，合规要至少要做到以下几点：

其一，采取合理措施实现数据匿名化。匿名化意味着，在符合比例原则的前提下，投入一定时间、成本努力也无法恢复身份属性。数据匿名化是一个相对概念。在可获得的数据源越来越广泛、数据算法越来越强大的形势下，匿名化的数据集存在可重新识别身份的可能性。为此，企业要对匿名的算法，关联关系予以保密。

其二，数据匿名化的商业利用项目，推荐进行安全风险评估。对于风险较高的项目，要采取配套限制措施。例如采取限制合作对象，限制利用方式等措施。

其三，在数据匿名化之后，商业利用应限定在非身份化的模式之下，在后续的利用中不得进行身份识别。如果在第一阶段完成了数据匿名化，但在后续利用阶段，又采取定位到具体个人的方式（如精准化营销），这将使得前一阶段的匿名化工作失去意义，整个项目又将重新落入到个人数据保护法的调整范围。

其四，如果企业将匿名化数据提供给其他第三方，需采取配套合规措施。企业必须与之签订合同，要求对方不能够将匿名化数据与其他信息进行比对、参照，以实现身份识别的功能，并采取 IT 审计等有效措施监督对方履行该项义务。