武漢大學學報(哲学社会科学版)
WUHAN UNIVERSITY JOURNAL (Philosophy & Social Science)
第77卷第4期 2024年7月
Vol. 77 No. 4 July 2024 128~140
DOI:10.14086/j.cnki.wujss.2024.04.012

个人信息范围的界定与要件判断

程 啸

摘 要 个人信息是个人信息保护法中最核心的概念,也是个人信息保护法律规范的适用前提。我国法律对个人信息采取了统一定义的模式。从《中华人民共和国网络安全法》(下文简称《网络安全法》)到《中华人民共和国个人信息保护法》(下文简称《个人信息保护法》),个人信息的概念经历了一个从窄到宽的发展演变过程。《个人信息保护法》第4条第1款将个人信息的判断要件分为积极要件与消极要件,前者是通过关联性与识别性去界定个人信息的概念范围,后者则将匿名化处理后的信息排除在个人信息之外。为了防止个人信息的范围过于宽泛,以至于个人信息保护法成为无所不包的法律,应明确个人信息积极要件中关联性要件与识别性要件之间为"且"的关系而非"或"的关系,以此来相应地控制个人信息的范围。认定关联性要件时,不仅应考虑信息的内容、目的和结果,还要考虑信息与个人权益是否存在一定的因果关联。在判断识别性时,应当限定识别主体的范围及所使用的手段与方式。作为消极要件的匿名化虽然并非可以绝对消除信息的可识别性,但是其对于保护个人信息权益也有重要的意义。

关键词 个人信息;个人信息保护法;关联性;识别性;匿名化
中图分类号 D913 **文献标识码** A **文章编号** 1672-7320(2024)04-0128-13
基金项目 清华大学自主科研计划文科专项经费项目(2021THZWYY02)

"个人信息"是个人信息保护法中最核心的概念,个人信息的认定及其范围问题对于任何国家的个人信息监管法律体系而言,都至关重要。只有当处理者所处理的信息是个人信息时,才应当适用个人信息保护的法律规范,处理者方负有各种法律义务(如告知并取得同意的义务、合规义务、保护个人信息安全义务等),个人针对此种处理其个人信息的活动才享有查阅、复制、更正、可携带、删除等权利。无论是行政机关履行个人信息保护监管职责,司法机关解决个人信息保护案件,还是企业等个人信息处理者的合规实践,都始终离不开对个人信息的认定。

随着网络信息科技尤其是大数据技术的高速发展,人类社会的信息愈发丰富,数据化程度越来越高,而现实与虚拟之间及人类、机器和自然之间的界限也在不断模糊。在数据爆炸性生成与聚合、数据分析突飞猛进的当下,几乎所有的信息都能够被合理地被认为是个人信息,由此就产生了个人信息的范围究竟应当顺势扩张还是需要适当限制的问题。从当今世界上对个人信息保护力度最大的欧盟来看,其不仅有着统一的、包含范围很宽的个人信息(也称个人数据)的概念,并且多年来欧洲法院及依据《数据保护指令(95/46/EC)》成立的第29条工作小组(下文简称"第29条工作小组")持续地对个人信息的要件进行扩张性解释。因此,欧盟法上个人信息的范围越来越广。由此导致一些学者担心,欧盟法上越来越宽泛的个人信息概念将会使个人信息保护法面临成为"无所不包之法"(the law of everything)的风险:一方面,其想要为所有的情形提供最高的法律保护;另一方面,现实中又根本做不到,最终导致个人信息保护法出现"系统超载"[1](P42)[2]。我国个人信息保护立法吸收借鉴了欧盟数据保护立法的不少有益成

果，《个人信息保护法》第4条第1款更是采取了与欧盟法大体相同的个人信息定义。随着网络信息社会与数字经济的飞速发展，我国同样面临着如何妥当界定个人信息范围的问题。

个人信息是一个抽象的法律概念。任何学科“在形成抽象概念时，其定义中需要选择哪些特征根本上取决于相关学科在形成概念时想要追求的目标”[3](P552)。我国《个人信息保护法》以保护个人信息权益与促进个人信息的合理利用为目标，因此，理解个人信息的概念也必须始终立足于这一立法目标。就我国个人信息的范围问题，本文的主要观点是，虽然我国借鉴了欧盟法中个人信息的概念，但考虑到我国的国情及个人信息保护法的立法目标，个人信息的范围不宜过度扩张。在我国《个人信息保护法》第4条第1款已经将“关联性”与“识别性”作为认定个人信息的积极要件，同时将“匿名化”作为排除个人信息的消极要件的前提下，应当通过谨慎合理的解释判断该三项要件来限定个人信息的范围，从而协调权益保护与信息自由的关系，以满足我国信息社会与数字经济发展的需要。本文将分以下五部分展开讨论：第一部分梳理我国法上个人信息概念从窄到宽的演变过程，第二部分分析我国《个人信息保护法》所确立的个人信息的要件及其相互间的关系，第三部分讨论控制个人信息范围的方法，第四和第五部分依次研究个人信息的积极与消极要件的判断。

一、个人信息概念的立法演进

总的来看，我国法上个人信息的概念经历了一个从窄到宽的演变过程，即从最初限于识别自然人的个人身份的信息，扩张为既包括身份识别信息也包括自然人的活动状态信息，再到识别自然人的各种信息，直至与已识别或可识别的自然人相关的各种信息。

（一）从识别“自然人个人身份”到识别“特定自然人”

2012年的《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1条第1款规定：“国家保护能够识别公民个人身份和涉及公民个人隐私的电子信息。”该规定尚未区分个人信息与隐私，但其关于能够识别公民个人身份的电子信息属于个人信息的规定被此后的《网络安全法》继受。2016年颁布的《网络安全法》第76条第5项规定：“个人信息，是指以电子或者其他方式记录的能够单独或者与其他信息结合识别自然人个人身份的各种信息，包括但不限于自然人的姓名、出生日期、身份证件号码、个人生物识别信息、住址、电话号码等。”该定义采取了识别说的定义模式，并将识别的对象限定于“自然人个人身份”。所谓自然人个人身份，主要是特定自然人的姓名、年龄、性别、身高、相貌、职业、工作单位、教育程度、家庭住址、电话号码等。按照这个定义，自然人的各种活动信息如行踪轨迹、网络浏览痕迹、交易消费活动等，只要不能直接或间接识别出自然人个人身份，就不属于个人信息。显然，这个范围过于狭窄，因为自然人的行踪轨迹等活动信息不仅属于个人信息，而且属于敏感的个人信息。故此，2017年《最高人民法院、最高人民检察院关于办理侵犯公民个人信息刑事案件适用法律若干问题的解释》对《网络安全法》中的“自然人个人身份”做了广义的理解。该司法解释第1条将《中华人民共和国刑法》（下文简称《刑法》）第253条之一规定的“公民个人信息”界定为“以电子或者其他方式记录的能够单独或者与其他信息结合识别特定自然人身份或者反映特定自然人活动情况的各种信息，包括姓名、身份证件号码、通信通讯联系方式、住址、账号密码、财产状况、行踪轨迹等”。也就是说，公民个人信息不仅包括公民个人身份识别信息（即能够识别出特定自然人身份的信息），也包括体现特定自然人活动的信息[4](P319)。

2021年施行的《中华人民共和国民法典》（下文简称《民法典》）第1034条第2款继续采取识别说的定义模式界定个人信息，但对个人信息的涵义与外延作了扩张。一方面，该款仍旧采取了概括+列举的定义方式，但列举的个人信息种类有所增加，即在《网络安全法》的基础上新增了“电子邮箱、健康信息、行踪信息”；另一方面，《民法典》第1034条第2款仅要求“能够单独或者与其他信息结合识别特定自然人”即属于个人信息，而不限于识别自然人个人身份或反映特定自然人活动情况。也就是说，只要这些信息

能够将自然人识别出来,就是个人信息。立法机关撰写的民法典释义书指出,构成个人信息的核心要件就是具有识别性,“所谓识别,就是通过该信息可以直接或者间接地将某一自然人‘认出来’”[5](P209)。

(二)从“识别性”要件到“识别性+相关性”要件

《个人信息保护法》的起草工作与《民法典》的编纂同时进行。《民法典》颁布后,不少人认为,《个人信息保护法》应当与《民法典》衔接,采取定义加列举的方式对个人信息作出界定。但是,立法机关并未将《民法典》第1034条第2款的定义照搬到《个人信息保护法》中,而是对个人信息做了一个新的定义。《个人信息保护法》第4条第1款规定:“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,不包括匿名化处理后的信息。”将该定义与《民法典》的定义相比较可知:首先,《个人信息保护法》没有再采取概括+列举的方式界定个人信息,它只界定了何为个人信息,未列举具体的个人信息。这样做的原因并非如某些学者所言,是因为个人信息类型的多样性、外延上的抽象性以及该法对敏感的个人信息作出了明确规定[6](P39),而是因为,列举个人信息的做法固然更形象直观,但容易令人误以为个人信息是固定不变的,误以为只要是法律列举的信息,无论何时何地都是个人信息。有些信息确实如此,如人脸、指纹、声纹、掌纹、基因、虹膜、耳廓等个人生物识别信息,因为是与特定的自然人唯一对应的且无法或很难改变的。但有些信息是否属于个人信息,必须要考虑到具体场景,如姓名、出生日期。以“张伟”这个姓名为例,放在全中国的范围,如不与其他信息相结合,显然是根本无法单独识别特定的自然人。仅2016年全国叫“张伟”的人就有299025人[7]。然而,如果在某小学一年级的某个班级中,“张伟”这个姓名可能足以识别特定的一个男生(或女生)。因此,《个人信息保护法》不去列举哪些信息,旨在强调的是个人信息的相对性[8](P22)。

其次,《个人信息保护法》在坚持个人信息认定标准中的“识别性”要件之外,还增加了“相关性”要件(也称“关联性”要件),即个人信息除了满足“识别性”要件外,还必须满足“关联性”要件,应当是与已识别或可识别的自然人有关的各种信息。如此一来,个人信息的范围在某种意义上又得到了扩展。一方面,如果某个信息处理者已经识别或能够识别特定的自然人,那么,对于该处理者而言,与该自然人有关联性的所有信息都是个人信息,即便某些信息对于其他的主体而言既无法单独识别也无法间接识别该自然人。例如,对于电信公司而言,客户都是实名的,因此,任何客户使用电信服务的时间、地点等信息对电信公司而言,都是个人信息。另一方面,即便某些信息不是关于某个自然人的,而是关于某物的信息,如房屋的价值信息、汽车的维修保养信息等,但由于该物是属于自然人或者被自然人使用的,因此,这些物的信息因为与自然人有关,也会成为个人信息。

二、个人信息的要件及其关系

我国《个人信息保护法》第4条第1款从积极和消极两个方面规定了个人信息的要件,其中,积极要件就是识别性与关联性,而消极要件就是将匿名化处理的信息排除在个人信息之外。

(一)积极要件与消极要件

1995年欧洲议会和欧盟理事会发布的《数据保护指令(95/46/EC)》第2条将个人数据界定为“是指与已识别或可识别的自然人(‘数据主体’)有关的任何信息;可识别的人是指可以直接或间接地加以识别,尤其是通过身份证号码或与其身体、生理、心理、经济、文化或社会身份相关的一个或多个因素加以识别”。2018年欧盟的《通用数据保护条例》总体上延续了这个定义,其第4条第1款规定:“‘个人数据’,是指与已识别或可识别的自然人(‘数据主体’)有关的任何信息。一个可识别的自然人是指能够被直接或间接地加以识别,尤其是通过诸如参考姓名、身份证号码、位置数据、在线标识符或者与该自然人的身体、生理、遗传、心理、经济、文化或社会身份相关的一个或多个因素加以识别的人。”欧盟学者的通说一般都是将欧盟法上的个人数据概念分解为以下四个组成部分:任何信息(any information);相关的(relating to);已识别或可识别(identified or identifiable);自然人(natural person)(Article 29 Data Protection

Working Party，Opinion 4/2007 on the Concept of Personal Data）[9]（P109）。

我国《个人信息保护法》第4条第1款吸收借鉴了《通用数据保护条例》的规定，其与欧盟法关于个人数据的界定基本相同，但也有所区别。《个人信息保护法》上个人信息的概念同样可以被分为四个组成部分：各种信息；有关的；已识别或可识别；自然人。这四个组成部分中，认定个人信息的要件就是第2项和第3项，即关联性与识别性，它们是个人信息的积极要件。然而，与欧盟不同的是，《个人信息保护法》直接将匿名化处理后的信息排除在个人信息之外，因此，匿名化成了个人信息认定的消极要件。欧盟法并未如此规定。《数据保护指令（95/46/EC）》序言部分的第26条指出，保护原则不适用于经匿名化处理以致数据主体不再可识别的数据。《通用数据保护条例》定义了何为匿名化，但未明确将匿名化处理的数据排除在个人数据之外，其序言部分第26条指出："数据保护原则不适用于匿名信息，即与已识别的或可识别的自然人无关的信息或者以使数据主体不再可识别的方式匿名化的个人数据。"

（二）关联性与识别性的关系

在我国《个人信息保护法》确立的认定个人信息的两项积极要件中，关联性与识别性之间究竟是什么关系？对此，有不同的认识。一种观点认为，识别性包含了关联性。个人信息的实质要素就是"识别"，而认定识别需要结合识别的判断基准，信息相关性和识别可能性等三方面加以判断，"相关要件"实质上蕴含在中国个人信息概念的"识别"要件之中，具有识别性则必然与自然人相关，不具有相关性必然无法识别自然人[10]（P43）。另一种观点认为，认定个人信息时，识别性和关联性只要满足其中之一即可。识别性是从信息到个人，关联性是从个人到信息。例如，国家推荐标准《信息安全技术 个人信息安全规范（GB/T 35273-2020）》的附录A指出："判定某项信息是否属于个人信息，应考虑以下两条路径：一是识别，即从信息到个人，由信息本身的特殊性识别出特定自然人，个人信息应有助于识别出特定个人。二是关联，即从个人到信息，如已知特定自然人，由该自然人在其活动中产生的信息（如个人位置信息、个人通话记录、个人浏览记录）即为个人信息。符合上述两种情形之一的信息，均应判定为个人信息。"

笔者认为，上述两种观点值得商榷。关联性与识别性是个人信息必须同时具备的要件，二者缺一不可。首先，关联性并非是包含在识别性中的。在信息并没有直接涉及特定自然人时，需要先分析其与个人是否具有关联，然后才需要考虑是否具有识别性。例如，通过联网的电冰箱收集到的冰箱开关门次数、耗电量等信息，先要考虑是否与自然人相关，满足关联性要件与否，然后分析该信息是否具有识别性。其次，当自然人已经被识别时，虽然不需分析识别性，却仍要考察是否满足关联性的要求。这是因为，并非所有的涉及已识别人的信息都是个人信息，如考官对于考生A作的答试卷的分析、政府针对个人B提出的不动产所有权转移登记申请作出决定前进行的内部法律评估分析等，这些信息中出现了特定的自然人A或B，但仍然需要考察该信息是否与个人相关，即满足关联性与否，否则就会不适当的扩张个人信息的范围（下文详述）。总之，关联性与识别性并非择一关系或包含关系，而是并列关系，二者共同发挥合理界定个人信息的范围的功能。

三、个人信息范围的控制方法

当今世界对个人信息保护的立法主要有两种：一是以欧盟为代表的统一立法模式，二是以美国为代表的分散立法模式。在个人信息保护统一立法模式中，个人信息或个人数据采取的是一种连贯的、统一的定义方式，如《通用数据保护条例》第4条第1款、我国《个人信息保护法》第4条第1款。在分散主义立法中，不存在对个人信息的一般性的界定，只有不同的法律对所调整的不同类型的个人信息的定义，例如，美国的联邦法律和州的法律没有统一的个人信息的定义，不同的法律采取的定义也各不相同①。这两种方式各有利弊。以美国为代表的个人信息分散定义的模式在认定个人信息时标准更具体、更具操

① 根据Schwartz与Solove教授的归纳，美国法上定义个人信息的方法主要就是三种，即"同义反复"法（tautological approach）、"非公开"法（non-public approach）以及"具体类型"法（specific-types approach）[11]。

作性,但弊端在于:其一,个人信息的范围比较狭窄;其二,不同法律中的个人信息的概念存在差异冲突;其三,给个人信息保护监管以及处理者的合规增加难度。以欧盟为代表的统一界定模式的好处在于,因为抽象,所以概念的适应性强,而缺点在于个人信息的范围难以被恰当的控制。如前所述,这种个人信息范围过度扩张的情形在欧盟法上已经出现。我国采取的也是统一定义模式,因此,如何解决我国法上个人信息范围的控制问题,值得研究。

(一)通过要件控制个人信息的范围

就如何避免我国法上个人信息范围的过度扩张,理论界提出了三种观点:第一种观点认为,应当限缩个人信息的概念,因为立法者保护个人信息的目的是在社会共识范围内确保自然人不愿意公开的私密空间、行为轨迹等相关信息处于私密状态,所以识别的对象应当是"特定自然人的身份或者行为轨迹"[12](P134)。第二种观点认为,应当取消个人信息与非个人信息的区分。个人信息高度依赖场景且因个人信息识别目的、识别主体、识别概率和识别风险的不同而不同,我国应当放弃个人信息与非个人信息的绝对区分,而对个人信息的概念采取场景化的界定方式,同时借鉴国外学者提出的"个人信息、可识别个人信息、非个人信息的三分法"作为指导我国个人信息保护监管的理论[13](P54-58)。第三种观点主张,在维持宽泛的个人信息的标准下,采取风险路径的方法将个人信息的识别性和相关性进行程度区分,依据相应的风险承担相应的保障义务[14](P74-75)。

笔者认为,上述观点值得商榷。首先,在《个人信息保护法》已经确立了范围更宽的个人信息的概念的前提下,重要的不是退回到以往的法律和司法解释采取的狭窄的个人信息概念,而是通过对个人信息构成要件的合理解释与判断来控制个人信息的范围。况且,将识别对象局限于特定的自然人的身份或行为轨迹的观点也混淆了个人信息权益与隐私权的差异,个人信息保护法并非只是保护私密信息,同样保护公开的个人信息。

其次,取消个人信息与非个人信息的区分观点并不妥当。个人信息是个人信息保护法的核心概念,如果被取消了,就无法划定个人信息保护法的适用范围,这将导致个人信息保护法的执法完全丧失可预期性。况且,所谓个人信息、可识别个人信息、非个人信息的三分法,仍然是要先区分个人信息与非个人信息,然后再将个人信息中又分出可识别的个人信息。所有的法律概念都具有一定的不确定性,个人信息的概念也不例外。所谓场景化的界定个人信息(尤其是敏感的个人信息)仍然是要通过个人信息的要件加以认定,无法在取消个人信息的概念后单独的实现场景化认定。

再次,《个人信息保护法》已将个人信息区分为敏感的个人信息与非敏感的个人信息,其中敏感的个人信息指信息被识别后对自然人会带来很高的风险,即第28条第1款规定的"一旦泄露或者非法使用,容易导致自然人的人格尊严受到侵害或者人身、财产安全受到危害的个人信息"[15](P85)。并且《个人信息保护法》对于敏感个人信息给予了更高程度的保护,如要求必须具有特定的目的和充分的必要性并采取严格保护措施的情形下,才可处理敏感的个人信息,还要求处理敏感的个人信息必须取得个人的单独同意等。因此,根据识别风险确立不同强度的义务的观点无法再用于控制个人信息范围扩张的问题。

总之,在我国《个人信息保护法》已经采取了与欧盟相同的宽泛的个人信息概念的大背景下,要控制个人信息的范围,关键就在于科学合理的解释与认定个人信息的要件。

(二)个人信息要件的解释与认定规则

在对个人信息的要件进行解释与认定时,应当遵循以下三项基本的规则。

1. 个人信息权益保护与个人信息合理利用的协调规则

个人信息的认定是个人信息保护法律规范适用的前提,某一信息被认定为个人信息,对信息主体而言,就享有了个人信息权益,从而能够行使针对个人信息处理者的查阅、复制、更正、补充、可携带、删除等各项权利。然而,个人信息权益意味着个人对其个人信息处理享有知情权和决定权,这就限制了个人信息的流动与利用。个人信息权益主体之外的人要利用这些个人信息,就必须有法律根据,并遵循合

法、正当、必要、目的限制等基本原则。一旦侵害个人信息权益，还需要承担法律责任。我国《个人信息保护法》第1条虽然将个人信息权益保护作为第一个立法目的加以规定，但也明确了促进个人信息的合理利用同样是一个立法目的。因此，在认定个人信息的积极要件或消极要件时，不能仅仅为了保护个人信息权益而做过度的扩张解释，这不仅会妨碍个人信息的合理利用，在我国法上还会给处理者造成很大的法律风险。这种风险不仅包括履行个人信息保护职责的部门施加的行政处罚，还意味着处理者可能需要承担刑事责任①。

2. 有利于降低个人信息保护法实施成本的规则

欧盟学者之所以担心个人信息的范围不断扩大的问题，倒不是因为他们认为个人信息的范围扩大是不对的，而是他们认为，扩大的个人信息的范围即意味着个人信息保护法适用范围的扩大，如果个人信息保护执法和司法跟不上，则个人信息保护法规定的权利义务和责任体系就无法落到实处，最终导致个人信息受保护的权利这一基本权利成为口号[1](P78-79)。我国同样如此。如果不适当的扩大个人信息的范围，不仅会导致个人信息处理者的合规成本直线上升，出现为了减少合规成本而非法处理个人信息的情形。此外，我国目前的个人信息保护执法力量本来就很薄弱，不仅履行个人信息保护职责的部门众多，九龙治水，难以统一，而且投入个人信息保护日常监管和执法的资源也很少。个人信息保护的监管工作当下还主要是通过间断性、运动性执法的方式加以完成。在既有的最典型的个人信息尤其是敏感的个人信息仍存在大量非法处理和非法买卖的乱象的情形下，监管机关根本无法应对那些被扩张进来的个人信息的保护问题。因此，在解释与判定个人信息的要件时，必须要考虑个人信息保护法的实施成本。

3. 个人信息保护法与其他法律相协调的规则

在万物互联的现代网络信息社会，只要是地球上的信息和数据都能从某种角度上牵扯到自然人，被认为是与已识别或可识别的自然人相关的信息。如果不做限制的话，就会导致所有人类社会的活动都要受到个人信息保护法的调整，这不仅极不适当地扩大了个人信息保护法的适用范围，还会造成个人信息保护法与其他法律的重叠与冲突。例如，只要各种法人或非法人组织的文件中出现了个人的姓名、身份证号等，就将这些文件中的信息都作为个人信息，个人就可以基于个人信息权益而针对处理者行使查阅、复制、更正、删除等权利，那么《个人信息保护法》就必然会与《中华人民共和国反不正当竞争法》《中华人民共和国保守国家秘密法》（下文简称《保守国家秘密法》）《中华人民共和国政府信息公开条例》（下文简称《政府信息公开条例》）等法律法规之间产生重叠与冲突。所以，解释与判断个人信息的要件时，应当注意协调个人信息保护法与其他法律的关系。

四、个人信息积极要件的判断

现代社会就是信息社会，但不是任何信息都属于个人信息，只有与个人有关联性的信息，并且该个人能够从中被识别出来的信息，方为个人信息。个人信息必须同时满足关联性与识别性这两个要件。

（一）关联性要件的判断

如何认定关联性，值得研究。目前比较具有代表性的判断标准有两个，一是欧盟法上三要素的标准，二是英国法院提出的接近性测试标准。

1. 三要素标准

第29条工作小组认为，判断某一个信息是否与某个自然人有关时，只需要满足“内容”“目的”或“结果”这三个元素中的一个即可。具体而言，当信息是关于某个个人的，即所涉信息的描述性内容是关于

① 我国《刑法》第253条之一规定了侵害公民个人信息罪，而司法实践是依据非常僵化的标准即不同类型的个人信息的数量来确定是否构成该罪的。参见《最高人民法院、最高人民检察院关于办理侵犯公民个人信息刑事案件适用法律若干问题的解释》第5条第3-5项的规定。

个人的,就满足了内容元素的要求。例如,医学分析报告的结果明显与特定患者有关,或者公司客户目录下的信息明显与该客户相关。如果信息在某一具体个案的各种情况下被用于或可能被用于评估、对待或影响某个人的行为或状态时,该信息就具备了目的要素,如公司办公室的电话使用记录。即便缺乏内容要素或目的要素,只要信息的使用会对某个人的权利或利益产生影响,即便该影响并非是主要的影响,但只要由于该信息的处理而使得该人相比于其他人可能会被差别对待,就满足了"结果要素"(Article 29 Data Protection Working Party, Opinion 4/2007 on the Concept of Personal Data)。

在2014年的"YS案"的判决中,欧洲法院并未采取第29条工作小组提出的上述三要素判断标准。该案中,三名第三国的公民(即"YS""M"和"S")申请在荷兰居留。M和S的申请获得批准,但YS的被拒绝。随后,三名申请人均要求获取荷兰当局有关其居留申请的内部文件。这些文件中包括申请人的姓名和种族等数据,还包括评估申请人申请的法律分析。在2009年7月14日之后,荷兰当局提供给申请人的副本中不再包含法律分析。申请人向荷兰法院起诉要求获得法律分析这部分信息,因为申请人想了解当局对其申请作出决定的理由,而这些理由包含在法律分析当中。该案中欧洲法院面临的关键问题就是,法律分析是否属于个人数据?欧洲法院在判决中对"关联性"要件进行了狭义的解释,其认为法律分析不属于个人数据,理由在于:其一,尽管法律分析部分可能包含了个人数据,但法律分析不是与居留许可的申请人有关的信息。在不限于对法律的纯粹抽象解释的情况下,该法律分析最多是关于主管机关对申请人情况的评估和法律适用的信息,而申请人的情况是通过该当局掌握的与其有关的个人数据确定的。其二,《数据保护指令(95/46/EC)》的目的是为了保护自然人的基本权利和自由,特别是保护隐私权,赋予申请人对法律分析这种信息以访问权等权利不符合《数据保护指令(95/46/EC)》保护隐私权的目的(YS v Minister voor Immigratie, Integratie en Asiel [2015] 1 WLR 609)。

然而,到了2017年的"Nowak案",欧洲法院又全面采纳了第29条工作小组提出的认定关联性的三要素说。在该案中,考生Nowak因为没有通过爱尔兰会计师协会组织的考试,故此,其基于个人数据访问权要求取得他的考卷的副本,被爱尔兰会计师协会拒绝,Nowak向数据保护官投诉,从而引发本案争议。欧洲法院认为,考生考试中的书面答案和考官对该答案的评阅意见属于《数据保护指令(95/46/EC)》中的个人数据,理由在于:首先,"个人数据"的概念潜在地包含任何信息,只要该等信息因其内容、目的或效果而与数据主体有关。就考生的答案和考官的评阅意见而言,它们在内容上反映了有关考生的信息(如知识水平,思维过程及手写答案时的笔迹信息,以及考官对考生表现的看法);从处理目的上看,是要评估考生的专业能力;从结果上看,使用这些信息"可能会影响他或她的利益",如确定考生的就业机会。其次,一旦信息被归类为个人数据时,那么整个数据保护法律体系都适用,包括施加给数据处理者的义务和赋予数据主体的访问、更正、异议、删除等权利。考生基于保护其私生活的合法利益,有权根据数据保护法行使针对所做答案和考官的评论进行访问、更正和异议权利。为考生提供这些权利符合数据保护法的目的(Peter Nowak v. Data Protection Commissioner, In Case C-434/16)。

2. 接近性测试

在2003年的"Durant案"中,英国上诉法院提出了一个"接近性测试(proximity test)"用来判断数据是否涉及"个人"。该案的原告在起诉巴克莱银行的诉讼中失败,金融服务管理局(FSA)随后对原告针对巴克莱银行的投诉进行了调查。原告要求FSA披露与他的投诉相关的信息,以便继续进行与巴克莱银行的纠纷。FSA披露了一部分信息,但拒绝披露另外一部分。该案的核心是,FSA拒绝披露的与其调查相关的信息是否属于英国《1998年数据保护法》中"个人数据"的含义。法官奥尔德在判决中指出:"不是从计算机中检索到的所有与个人姓名或唯一标识符有关的信息都是《1998年数据保护法》中的个人数据。数据控制者所持有的文件中仅仅提及了数据主体,并不一定就构成该数据主体的个人数据。在任何特定情况下,是否构成个人数据取决于其在与数据主体连续关联或连续接近中所处的位置,而不是或多或少的涉及数据主体参与的交易或事件。在我看来,有两个概念可能有所帮助。首先,信息是否在典

型的意义上属于传记性的，即超越了对假定为数据主体的个人参与的不具有个人内容（即其隐私不能被认为受到损害的生活事件）的事项或事件的记录范围。第二个是信息的焦点。信息应以假定的数据主体作为焦点，而不是以他可能参与的或感兴趣某些交易或事件中涉及的其他人为焦点，例如，在本案中对于可能经过他唆使的某些其他人或机构的行为的调查。简言之，该信息是影响他的隐私的信息，无论是与他的个人生活、家庭生活还是在商业、专业能力等方面有关。”（Durant v Financial Services Authority [2003] All ER (D) 124 (Dec)）通过“Durant案”，英国上诉法院建立了所谓的“接近性测试”的概念。据此，有学者认为，在认定个人信息的关联性要件时，应当在第29条工作小组提出的“内容”“目的”和“结果”三要素判断标准的基础上，再进行接近性测试。在接近性测试中，信息只有在与个人足够接近时才是与个人相关的，即涉及个人的。这种测试主要用于评估所涉信息的内容（即评估所涉信息在多大程度上描述了作为数据主体的个人而不是其他内容），但也能适用于评估信息使用的目的或结果与个人是否具有关联性。不过，此时应当遵循相对严格的标准，即信息应当对于所涉及的个人具有很高程度的具体性[2]。

3. 我国应采取的标准

就我国应采取何种关联性要件的判断标准问题，有学者主张直接采用第29条工作小组提出的三要素判断标准[16]（P14-15），还有的学者认为应当采用宽泛的定义，即只要存在用于分析个人特征的可能性就属于个人信息[14]（P74）。笔者认为，第29条工作小组的对“关联性”的认定过于宽泛，这使得“关联性”要件对于个人信息的范围几乎起不到任何的控制作用，与个人存在任何联系（无论该联系多么间接和遥远）的信息都可能被认为是个人信息。例如，一份文件中仅仅因为出现了某个自然人的姓名，那么就具有关联性，加之能够从该文件中识别出特定的自然人，于是全部文件就都成了个人信息[2]。如前所述，这样做的结果虽然有利于加强个人信息保护，但也会导致社会生活中大量的信息处理活动被纳入到个人信息保护法的调整范围，使处理者因此而负担作为个人信息处理者的法律义务与责任，极大地增加政府机关的执法成本以及企业的合规成本，还导致了个人信息保护法与其他法律的冲突。我国《个人信息保护法》之所以确立个人信息权益并严格加以保护，不是为了保护个人信息本身，也不是要确立自然人对其个人信息这一客体享有如同所有权那样的排他的支配的效力，而只是通过确立自然人对其个人信息的处理享有知情权和决定权，以实现防止与避免自然人因其个人信息被非法处理而致人身财产权益遭受侵害或人格尊严、个人自由受到损害的目的[17]（P37）。因此，我国在认定个人信息的关联性要件时，应当综合吸收借鉴三要素标准与接近性测试的合理因素，将二者加以结合。具体而言，首先，如果某个信息就是以可能属于信息主体的某个自然人为中心内容的，是关于该自然人的个人身份、特征或行为等所谓传记性内容的，那么，该信息就与自然人具有足够的接近性，应当认为是与自然人有关的信息。比较典型的如医疗档案关于病人的病情、学习档案关于某个学生的学习信息等，就是专门针对这个病人或者学生的。如果信息的内容是关于某个事件、事项或交易的记录，而并非是对作为可能的信息主体的个人的身份、特征或行为模式等内容的记录，即便该自然人也是被记录的事件、事项或交易的参与人之一，那么信息也不是与个人具有关联性的信息。

其次，如果某个信息的内容并没有直接涉及可能成为信息主体的自然人，只是关于某物的价值、使用、交易等信息，但是该信息在诉争案件的情形下能够可以被合理的预见对其之处理可能会对信息主体的个人权益产生不利影响的，也应当认为该信息满足关联性。例如，一辆汽车的行驶里程、油耗、维修等数据只是关于该汽车本身的数据，并不直接涉及个人，但是如果在诉争案件中对这些数据的分析就是为了预测司机的行为模式、心理性格，由于该数据的处理结果意味着对于司机行为的监视与操控，有害于人格尊严与人格自由。因此，该信息属于个人相关的信息。所谓“个人权益”不限于隐私权和个人信息权益，它是我国《个人信息保护法》第6、8、24、27、30、51条等条文中规定的“个人权益”，即自然人的各种受法律保护的权益，既包括宪法上的基本权利如人格尊严和人身自由，也包括《民法典》规定的自然人的

人身权益、财产权益,以及《中华人民共和国未成年人保护法》等法律特别规定的某类自然人享有的民事权益。

最后,如果某个信息不满关联性要件,不与活着的自然人相关,就无需考察其是否具有识别性。该信息不属于个人信息,自然人无权通过行使个人信息权益来对该信息进行查阅、复制、更正、补充、删除等。就该信息的保管、公开、查阅、利用等,应当适用相应的法律法规如《保守国家秘密法》《中华人民共和国档案法》《政府信息公开条例》等的具体规定。

(二)识别性要件的判断

识别性是判断个人信息的另一个要件。《民法典》将具有识别性的信息分为两类:一是具有直接识别性的信息,即只要通过该信息而无需借助其他信息就可以直接识别某一个自然人的信息,如自然人的身份证号码、基因信息等;二是具有间接识别性的信息,即仅仅通过该信息并不能直接识别自然人,而必须与其他信息相互结合后才能识别特定的自然人,如自然人的姓名[5](P209)[18](P430)。《个人信息保护法》第 4 条第 1 款没有再从信息本身是如何识别自然人的角度作出规定,而是要求自然人必须是"已识别或者可识别的自然人"。所谓"已识别(identified)"的自然人就是指特定的自然人已经被识别出来,至于该自然人是通过某一个信息就被识别出来的,还是多个信息结合后才被识别的,在所不问。"可识别(identifiable)"就是指特定的自然人具有被识别的可能性或概率。就识别性要件的判断而言,需要研究的是识别的可能性。

判断识别的可能性,需要解决两方面的问题:一方面,识别的主体是谁,即识别的可能性究竟是对谁来说的?是对特定的信息处理者,还是任何人?另一方面,识别的手段如何,即在判断识别可能性时,识别主体采取的手段和方法究竟要到何种程度?判断识别可能性有两种方法:一是绝对的方法(absolute approach),该方法意味着在判断信息能否识别自然人时要考虑处理者及其他任何人能够采取的所有方法和手段,无需考虑成本与费用的问题,也就是说,一切能够识别信息主体的可能性和机会都必须加以考虑。例如,就加密的数据而言,只要世界上有人能够解密,该数据就是个人数据。二是相对的方法(relative approach),即只考虑特定的信息处理者为了识别自然人而需要付出的必要努力,只有存在识别自然人的实际机会的情况下,才认为信息是可识别的,不会把那种纯粹理论上的识别风险作为识别的可能性[19](P165-166)。

就识别可能性的认定,欧盟法于识别主体的问题上采取了绝对标准,而在识别手段的问题又采取了相对标准。《数据保护指令(95/46/EC)》序言第 26 条指出:"为了确定某个个人是否可识别,应当考虑控制者或任何其他的人可能合理使用的全部的用于识别该个人的手段。"在 2016 年的"Breyer 案"中,欧洲法院认为识别的主体不限于控制者,还包括其他人。该案原告 Breyer 先生访问了一些德国联邦机构的公开网站,后者为了防止网络攻击而存储了访问者 Breyer 先生的动态 IP 地址。Breyer 认为,动态 IP 地址是他的个人数据,反对德国联邦机构收集该数据。静态 IP 地址是固定不变的,允许连续识别。但动态 IP 地址有所不同,其在每次建立新的网络连接时都会变化,并且也不允许通过公开的文件在特定计算机和互联网连接之间建立联系。因此,德国联邦机构无法单独通过动态 IP 地址来识别 Breyer 的,必须还要获取网络服务提供商提供的额外信息才能做到。由此,欧洲法院面临的核心问题是:此种动态 IP 地址是否属于可识别自然人的信息?欧洲法院认为:首先,个人数据并不要求其本身就能单独识别数据主体,无论单独识别还是与其他数据结合起来才能识别,都是个人数据。其次,使得个人被识别的所有信息并不是必须要由一个人拥有,在认定识别性时需要考虑到的所有可能合理使用的手段是由控制者还是由任何其他人掌握,无关紧要。再次,在考虑联邦机构掌握的动态 IP 地址与网络服务提供商持有的附加数据结合起来是否构成一种可能合理用于识别数据主体的手段时,需要考虑其可能性大小。如果是法律所禁止的或者因为要付出不成比例的时间、成本和人力以至于实际上是不可能时,那么将动态 IP 地址与附加数据结合的可能性就基本上没有(Patrick Breyer v.Bundesrepublik Deutschland, In Case C-582/14)。

就合理使用的识别手段问题，第29条工作小组认为，这意味着仅仅有可能单独识别个人的假设的可能性并不足以将该人视为“可识别”的自然人。如果考虑“控制者或任何其他人合理可能使用的所有手段”，这种可能性不存在或微不足道，则该人不应被视为“可识别”，信息也不会被视为“个人数据”。“控制者或任何其他人合理可能使用的所有手段”这一标准应特别考虑所有相关因素。识别的成本是一个重要的因素，但不是唯一的因素。在判断识别的合理可能时一方面需要考虑所有的相关因素，另一方面对于识别可能性的测试也是一个动态的过程。具体而言，需要考虑的因素包括但不限于：(1)识别的成本。(2)预期的明示或默示的处理目的(当信息的处理仅在允许识别特定个人并以某种方式对待他们时才有意义的话，应当假定识别工具以合理的可能存在)。(3)组织功能失调(如违反保密义务)和技术故障的风险，包括数据泄露。(4)信息处理时的技术，包括在处理的生命周期里面未来技术的可能发展。例如，今天可能无法通过所有合理可能使用的手段进行识别，并且由于数据的存储很短，也就是一个月，不大可能在该信息的生命周期内进行识别，因此不应将其视为个人数据。但是，如果存储期限是十年，则控制者应考虑可能在信息存储的第九年发生的识别可能性，而在那一时刻该信息就成了个人数据。(5)防止数据识别的措施(即保持匿名)很重要，这是作为一种避免完全处理个人数据的手段，而不是履行《数据保护指令(95/46/EC)》所规定的数据安全义务(Article 29 Data Protection Working Party，Opinion 4/2007 on the Concept of Personal Data)。欧盟《通用数据保护条例》吸收了《数据保护指令(95/46/EC)》与第29条工作小组的意见，其序言部分第26条指出：“为判断自然人身份是否可识别，需要考虑所有可能使用的手段，比如利用控制者或其他人来直接或间接的确认自然人身份。为判断所使用的手段是否可能用于识别自然人，需要考虑所有客观因素，包括对身份进行确认需要花费的金钱和时间，考虑现有处理技术以及科技发展。”

我国法没有对识别的主体究竟是谁以及如何认定识别手段等作出规定。司法实践中法院在认定识别可能性常常考虑的因素包括：信息的处理场景、处理方式、信息处理主体对于信息的控制范围和能力、识别的成本(如技术门槛、第三方数据来源、经济成本、还原时间)等①。笔者认为，就识别的主体问题，由于不同识别主体的技术能力、掌握的信息存在很大的差别，如果以所有的人是否具有识别的可能，即便如欧盟那样限制为合理使用的手段来判断，个人信息的范围也会变得非常大。个人信息的处理都是在具体的场景为了特定的目的以特定方式进行的，如果完全超越信息处理者以任何人作为识别主体来判断识别的可能性，显然是不合理的。

比较妥当的做法是：首先，识别主体原则上限制于诉争案件中的信息的处理者，但是，当信息处理者与其他掌握辅助信息的主体(即其信息与处理者的信息结合后能识别自然人的信息)之间存在法律或经济上的密切联系，使得在认定识别性时必须将该主体也纳入进来情形下，就应当以处理者和关联方作为识别主体。例如，处理者与A公司是同一集团的公司，是母子公司或者存在控股关系；再如，收集动态IP地址的处理者依据法律规定有权直接从B通信公司调取其他辅助信息从而识别特定自然人的，那么就应当将A公司与B通信公司也作为识别的主体加以考虑。

其次，对识别可能性的判断正如现代社会对“风险”的界定一样，都不是一个纯粹自然科学的问题。诚如著名社会学家乌尔里希·贝克所言，在界定何为风险、什么是可接受的水平时，不再是科学理性占主导地位，而是存在科学理性与社会理性的冲突，总是存在各种现代性主体与受影响群体的竞争性和冲突性的要求、利益和立场，它们被迫以原因和结果、策动者和受害者的方式一起去界定风险。风险界定在本质上就是“利益的博弈”[20](P28-31)。同样，可能性的认定也是协调信息自由与权益保护这两种冲突利益的结果，合理标准是相当广泛的且依赖于具体场景的，因此，信息作为个人信息的地位也是动态的、

① 比如“罗某与北京大生知行科技有限公司网络侵权责任纠纷案”(北京互联网法院[2021]京0491民初5094号民事判决书)、“余某与北京酷车易美网络科技有限公司隐私权纠纷案”(广州互联网法院[2021]粤0192民初928号民事判决书)。

变化的,构成个人数据的信息清单是会随着技术变化而发展的[21](P57)。总之,在识别的方法上,需要考虑识别成本、处理目的、处理方式、技术的发展状况、个人信息泄露的风险、存储期限等因素来进行动态的认定。

五、个人信息消极要件的判断

匿名化(Anonymisation)是指个人信息经处理无法识别特定自然人且不能复原的过程(《个人信息保护法》第72条第4项)。由此可见,匿名化是一种修改个人信息的方法,其结果是使得信息与个人没有关联。匿名化处理后的信息必须是无法识别特定自然人并且不能复原。对个人信息进行匿名化处理往往是在统计和科学研究中,其主要的方法有随机化(Randomization)与泛化(generalization)。匿名化不同于加密、去标识化,依据《个人信息保护法》第51条,后两者是个人信息处理者依法采取的安全技术措施,其中,去标识化(de-identification)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(《个人信息保护法》第72条第3项)。加密一般是在数据存储和传输中运用,通过加密算法将信息编码为残缺不全的状态,从而使得未经授权的人无法读取信息,只有获得密钥的人才能读取。经过加密后的个人信息可以很好地防止被他人未经授权的访问或被非法窃取或篡改。匿名化与去标识化、加密都有利于保护个人信息和隐私,可以降低信息主体的风险,也有利于处理者更好地履行义务。它们的区别在于:匿名化处理的信息不是个人信息,不适用个人信息保护法的规定。加密的个人信息以及去标识化的个人信息仍然是个人信息①。

(一)匿名化处理的信息不是个人信息

我国早在《网络安全法》中就对匿名化作出了规定,但只是将匿名化处理作为向他人提供信息可以无须被收集者同意的例外情形,即第42条第1款规定的"经过处理无法识别特定个人且不能复原的除外"。《个人信息保护法》第4条第1款界定个人信息时,首次明确将"匿名化处理后的信息"排除在个人信息之外。在审议个人信息保护法草案时,有些常委和专家对此有不同的看法。他们认为,匿名化处理固然可以大幅度提高还原个人信息的难度和成本,但是通过一定的手段仍有可能识别个人身份,且在现实中不排除一些企业以匿名化为由滥用个人信息,故此我国《个人信息保护法》不应当将匿名化处理的信息排除在个人信息之外。不过,最终立法机关"根据我国大数据发展和应用的实际,借鉴一些国家和地区的做法,明确经匿名化处理的信息不属于个人信息"[8](P23)。

笔者认为,将匿名化处理的信息排除在个人信息之外是有必要的。匿名化处理后的信息不属于个人信息,自然就不适用个人信息保护的法律规范,处理者免除了大量的义务,也不会因为违反个人信息保护法而承担民事、行政乃至刑事法律责任,有利于促进信息的流动和利用。在我国正在大力构建数据产权制度的当下,匿名化更是可以很好地有助于发挥"保护个人信息权益,促进数据流通利用"的重要作用。正因如此,《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》才明确提出要创新技术手段,推动个人信息匿名化处理,保障使用个人信息数据时的信息安全和个人隐私。

(二)匿名化处理后的再识别

个人信息的匿名化的要求是个人信息经处理无法识别特定自然人且不能复原。问题是,一方面,信息经处理后是否无法识别特定自然人,这本身就是相对的;另一方面,能否复原,也不是绝对的。在可获得的数据来源越来越丰富、数据量越来越大以及算法、数据分析越来越强的现代社会,匿名化处理后的信息依然存在或大或小的重新识别特定自然人的可能性。实践中已有不少事例证明了此点。例如,2000年美国计算机专业的教授Latanya Sweeney进行的一项研究表明,只要通过美国的邮政编码(ZIP

① 有的学者认为,如果采取的是最高程度的加密技术,信息处理者没有密钥或其他解密方法时,这就意味着处理者通过合理的手段是无法读取个人信息的,因此,该加密的个人信息就是匿名化处理的信息,应当视为非个人信息[21]。

码）、出生日期和性别这三个信息的组合，就足以识别美国87%的人口[22]（P2）。2008年，人们以公开获取的互联网电影数据库为基础，重新识别了50万奈飞（Netflix）用户的电影评级记录。再如，2013年，根据“匿名化”处理的纽约出租车乘车记录公共数据库以及狗仔队的照片，可以推断出如布莱德利·库珀（Bradley Cooper）和奥立薇娅·玛恩（Olivia Munn）等名人的行踪路线，包括街道地址以及他们是否付过小费[1]（P47）。故此，一些学者认为，随着大数据、云计算等新技术的兴起，传统的仅仅删除姓名和社保号码的匿名化技术已经失败了，技术专家可以通过再识别（re-identify）或者去匿名化（de-anonymize）的方法来实现个人身份的再识别[23]（P1716）。总之，匿名化不是绝对的，只是相对的，是在特定的时空和技术背景下来认定的。随着技术的发展和普及如云计算、量子计算机等，以及可能出现的其他附加信息的泄露，匿名化的信息被还原成为个人信息的可能性会逐渐发生变化。

尽管匿名化后的信息是否真的就不能识别自然人以及能否复原，都只是相对而言。但必须注意的是，无论如何匿名化技术都在相当程度上降低了信息被再识别的可能性，对于保护自然人的个人信息权益、隐私权都是有益的。因此，司法实践重要的问题在于：如何认定处理者提出的信息被匿名化处理后不属于个人信息这一抗辩。对此，应注意以下几点：其一，要区分被匿名处理的个人信息的类型。对于敏感的个人信息，匿名化处理的要求应当更高，即再识别的难度应当大大高于非敏感的个人信息。在非敏感的个人信息中又要区分公开的个人信息与非公开的个人信息。对于公开的个人信息，处理者本来就可以不经信息主体同意而在合理范围内处理，故此，对于匿名化处理的强度无需要求太高。其二，考虑匿名化处理后的信息的用途，不同的用途意味着被匿名化处理的信息与其他信息进行关联而重新变得具有识别可能的概率也是不同的。其三，匿名化的技术类型以及技术未来的发展状况，不同的匿名化技术产生的在识别的风险是不同的，而且会随着时间的推移发生变化。其四，复原的成本，具体包括所需的辅助信息获取的难易程度、所需花费的时间和金钱等。

在我国《个人信息保护法》已经采取了统一且宽泛的个人信息概念的前提下，面对客观上产生的个人信息的范围不断扩大的趋势，正确的解释与判断关联性、识别性以及匿名化等个人信息的要件非常重要。唯其如此，方能合理的控制个人信息的范围，有效地实现个人信息权益保护与信息自由流动利用，个人信息保护法与其他法律之间关系的协调，既避免个人信息保护的制度供给不足，又防止个人信息保护法承受不能承受之重。

（东南大学法学院王苑老师对本文初稿提出了宝贵意见）

参考文献

[1] Nadezhda Purtova. The Law of Everything: Broad Concept of Personal Data and Future of EU Data Protection Law. *Law Innovation & Technology,* 2018, 10(1).

[2] Benjamin Wong. Delimiting the Concept of Personal Data after the GDPR. Legla Studies. *The Journal of the Society of Legal Scholars*, 2019, 39(3).

[3] 卡尔·拉伦茨.法学方法论：全本.黄家镇译.北京：商务印书馆，2020.

[4] 喻海松.网络犯罪二十讲.北京：法律出版社，2022.

[5] 黄薇.中华人民共和国民法典人格权编解读.北京：中国法制出版社，2020.

[6] 江必新，郭峰.《中华人民共和国个人信息保护法》条文理解与适用.北京：人民法院出版社，2021.

[7] 张文.2016大数据“看”中国父母最爱给宝宝起什么名.人民日报（海外版），2017-01-10.

[8] 杨合庆.中华人民共和国个人信息保护法释义.北京：法律出版社，2022.

[9] Christopher Kuner, Lee A. Bygrave, Christopher Docksey. *The EU General Data Protection Regulation (GDPR): A Commentary.* Oxford: Oxford University Press, 2020.

[10] 韩旭至.个人信息的法律界定及类型化研究.北京：法律出版社，2018.

[11] Paul M. Schwartz, Daniel J. Solove. The PII Problem: Privacy and a New Concept of Personally Identifiable Information.

New York University Law Review, 2011, 86(6).
[12] 赵精武.个人信息"可识别"标准的适用困局与理论矫正——以二手车车况信息为例.社会科学,2021,(12).
[13] 丁晓东.论个人信息概念的不确定性及其法律应对.比较法研究,2022,(5).
[14] 谢琳.大数据时代个人信息边界的界定.学术研究,2019,(3).
[15] 王苑.敏感个人信息的概念界定与要素判断——以《个人信息保护法》第28条为中心.环球法律评论,2022,(2).
[16] 江必新,李占国.中华人民共和国个人信息保护法条文解读与法律适用.北京:中国法制出版社,2021.
[17] 程啸.民法典编纂视野下的个人信息保护.中国法学,2019,(4).
[18] 王利明,程啸.中国民法典释评·人格权编.北京:中国人民大学出版社,2020.
[19] Gerald Spindler, Philipp Schmechel. Personal Data and Encryption in the European General Data Protection Regulation. *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*, 2016, 7(2).
[20] 乌尔里希·贝克.风险社会:新的现代性之路.张文杰、何博闻译.北京:译林出版社,2018.
[21] 马里厄斯·克里奇斯托弗克.欧盟个人数据保护制度:《一般数据保护条例》.张韬略译.北京:商务印书馆,2023.
[22] L. Sweeney. Simple Demographics Often Identify People Uniquely. *Carnegie Mellon University: Data Privacy Working Paper*, 2000.
[23] Paul Ohm. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *University of California Los Angeles Law Review*, 2010, 57(6).

Definition of the Scope of Personal Information And Judgment of the Elements

Cheng Xiao（Tsinghua University）

Abstract　Personal information is the core concept in the Personal Information Protection Law and serves as the prerequisite for the application of all legal norms of personal information protection. Chinese laws have adopted a unified definition model for personal information, and the scope of the concept has been gradually broadened from *The Cybersecurity Law of the People's Republic of China* (hereafter *The Cybersecurity Law*) to *The Personal Information Protection Law of the People's Republic of China* (hereafter *The Personal Information Protection Law*). Paragraph 1, Article 4 of *The Personal Information Protection Law* divides the elements for judging personal information into affirmative and negative ones, with the former defining the scope of personal information through relevance and identifiability, while the latter excluding anonymized information from personal information. In order to prevent the scope of personal information from being too broad and making the Personal Information Protection Law all-encompassing, it is necessary to specify that the relevance and identifiability elements for personal information are connected by "and" rather than "or", thereby keeping the scope of personal information in check accordingly. To determine relevance, it needs to consider the content, purpose, and result of the information, and whether there is a causal relationship between the information and individual rights. To assess identifiability, the scope of the identifiable subject and the methods and means to be used should be restricted. While the negative element of anonymization may not eradicate the identifiability of information, it is still significant for protecting the rights and interests of personal information.

Key words　personal information; personal information protection law; relevance; identifiability; anonymization

■ 作者简介　程　啸,清华大学法学院教授,北京 100084。
■ 责任编辑　李　媛