doi:10.11835/j.issn.1008-5831.2018.02.011

欢迎按以下格式引用:齐爱民,张哲.识别与再识别:个人信息的概念界定与立法选择[J].重庆大学学报(社会科学版),2018(2):119-131.

Citation Format: QI Aimin, ZHANG Zhe. Identification and reidentification: The definition of personal information and the legislative choice[J]. Journal of Chongqing University(Social Science Edition),2018(2):119-131.

识别与再识别:个人信息的概念界定与立法选择

齐爱民,张 哲

(重庆大学 法学院,重庆 400044)

摘要:大数据时代,数据已成为社会生产和经济发展的关键要素,个人信息在提高政府决策水平、企业精准营销、社会管理创新等方面具有巨大的潜在利用价值。纵观全球,以欧盟为代表的个人信息保护单行立法模式成为当前世界各国的主流做法。国际社会在个人信息的界定上基本形成了以可识别性为核心判定标准的共识;但个人信息界定的动态性和场景性不仅带来了司法认定上的困难,也使企业在匿名化处理问题上无所适从。充分借鉴国外立法,以加强个人信息权顶层设计为核心,通过事前同意、事中风险评估和事后个案认定机制来弥补个人信息界定的固有缺陷,是提升中国未来民法典人格权编和个人信息保护法科学性的应有之义。

关键词:个人信息;再识别;匿名化;个人信息权

中图分类号:D923.4 **文献标志码**:A **文章编号**:1008-5831(2018)02-0119-13

信息社会中,随着全球网络基础设施的不断完善,移动互联网、物联网、云计算服务的普及,世界数据总量迎来了爆炸式增长。据 IDC 报告,未来全球数据总量年增长率将维持在 50% 左右,到 2020 年,全球数据总量将达到 40ZB。其中,中国数据量将达到 8.6ZB,占全球的 21% 左右[1]。在这其中,大部分是与互联网用户有关的个人信息。与此同时,以谷歌、微软、阿里、腾讯为代表的互联网企业在数据挖掘、人脸识别、机器学习等领域的技术也日臻成熟,大大提高了网络服务的个性化和智能化。

数据分析技术的进步和数据量的增长在便捷生活、创造经济价值的同时,也造成公民隐私的担

修回日期:2017-12-18

基金项目:国家社会科学基金重大项目"国家网络空间安全法律保障机制研究"(13&ZD181)

作者简介:齐爱民(1970—),男,河北晋州人,法学博士,重庆大学法学院教授,博士研究生导师,广西民族大学广西知识产权发展研究院院长,主要从事网络信息法与知识产权法研究,Email:1309978103@qq.com;张哲(1992—),男,河南平顶山人,重庆大学法学院博士研究生,主要从事信息法、知识产权法研究。

忧。美国以隐私权为核心的保护模式已无法抵御现代信息技术对私密生活的侵袭,于是有学者提出了“信息隐私法(Information Privacy Law)”和“个人可识别信息(Personal Identifiable Information)”等概念来进一步扩张隐私权的内涵。欧盟也通过颁布《一般数据保护条例(GDPR)》(以下简称“GDPR”)来进一步规制数据控制者和处理者的数据处理行为,以在促进信息自由流动的同时增强个人的数据控制力。纵观世界范围内的个人信息保护立法,虽然可识别性标准已经得到了绝大多数国家的认可,但在具体的外延界定上仍存在一定的差异。作为个人信息保护法中最为核心的概念,对其进行学理上的界定对于立法、司法和执法都至关重要,过窄的范围无法充分保护信息主体的合法利益,而过宽的范围也将阻碍信息的自由流动。以当前的社会发展状况为背景,通过对现有个人信息界定模式的反思,提出相应的解决对策对于中国未来个人信息保护法的制定具有重要意义。

一、个人信息概念界定现状

所谓个人信息,是指自然人的姓名、出生年月日、身份证号码、户籍、遗传特征、指纹、婚姻、家庭、教育、职业、健康、病历、财务情况、社会活动,以及其他可以识别该个人的信息[2]。纵观世界范围内的个人信息保护立法,主要存在两种模式。欧盟等采取统一立法的国家和地区普遍采用了“个人数据(PersonalData)”的概念,其是指与一个身份已识别或可识别的自然人(数据主体)相关的任何信息①。美国等采取分散式立法的国家则使用“个人可识别信息(PII)”的概念。从概念上看,两大法系在个人信息内涵的界定上呈现出趋同的倾向,但其外延却并不相同。为此,有必要对当前个人信息界定模式和外延予以梳理,以期为中国个人信息概念的确定提供借鉴和参考。

(一)个人信息界定模式

据不完全统计,当前世界上拥有个人信息保护法的国家(地区)近 90 个[3],它们在个人信息界定模式上主要体现为两种:一是以美国《儿童在线隐私保护法(COPPA)》(以下简称“COPPA”)、欧盟 GDPR、中国台湾“个人资料保护有关规定”、中国澳门《个人资料保护法》为代表的定义加列举的方式;另一种为单纯定义方式,如英国《数据保护法案(DPA)》(以下简称“DPA”)、法国《信息、档案与自由法》、德国《联邦数据保护法(BDSG)》(以下简称“BDSG”)、新加坡《个人数据保护法》,以及中国香港《个人资料(私隐)条例》均采用了此种界定模式。除此之外,相关国际组织亦在其官方文件中对个人信息进行了界定,如经济合作组织(OECD)在其《OECD 隐私框架》中就指出,个人数据是指任何与一个已识别或可识别的个人(数据主体)相关的信息②。亚太经合组织(APEC)发布的《APEC 隐私框架》认为,个人信息是指任何与一个已识别或可识别的个人相关的信息③。上述国家(地区)在个人信息界定模式上均采取了单纯定义的方式,没有明确肯定或者将特定信息排除在个人信息保护范围之外。

---

①Data Protection Directive. Article 2.

②Organisation for Economic Co – operation and Development. Guidelines governing the Protection of Privacy and Transborder Flows of Personal Data (2013). Part one.

③Asia – Pacific Economic Cooperation. APEC Privacy Framework. part ii. Scope. 9. Personal information.

可见,采用单纯界定的方式居多,立法者仅仅指出构成个人信息的核心要素,如自然人、可识别性等,而没有明确将姓名、身份证号、电话号码、社会安全码、基因信息等列为个人信息,也没有将特定的信息排除在外,从而保持了概念的开放性。

据统计,中国目前的个人信息保护规范已近 100 部,在规范制定主体上,除了全国人大常委会外,还包括国务院有关部委、地方政府、行业协会、科研机构等;在规制范围上则涵盖了互联网、电信、征信、证券、银行、保险、医疗等行业,可谓范围广泛,内容繁杂。尽管如此,多数学者仍指出中国个人信息保护规范存在着碎片化、保护利益不清晰、效力层级低、执法部门定位和权限不明确等问题[4],[5]33。纵观中国现行规范,同样存在单纯定义和定义加列举两种模式。

(1)定义加列举模式。工信部《电信和互联网用户个人信息保护规定》《电信和互联网服务用户个人信息保护定义和分类》、中国科学技术法学会与北京大学互联网法律中心《互联网企业个人信息保护测评标准》、中国互联网协会《互联网终端安全服务自律公约》《互联网终端软件服务行业自律公约》、中国广告协会互动网络分会《中国互联网定向广告用户信息保护框架标准释义和基本指引》等均采取此种界定模式,除了给出个人信息的一般定义外,还列举典型的个人信息类型以及排除类型。

(2)单纯定义模式。工信部《规范互联网信息服务市场秩序若干规定》、国家质检总局与国家标准委《信息安全技术公共及商用服务信息系统个人信息保护指南》、中国广告协会互动网络分会《互联网定向广告个人信息保护声明》、国家质检总局与国家标准委《健康信息学推动个人健康信息跨国流动的数据保护指南》、中国人民银行《中国金融移动支付检测规范第 8 部分:个人信息保护》等均使用了单纯定义的方式,并未列举典型的个人信息类型。从现行规范的定义模式看,采用单纯定义和定义加列举的做法数量相当,并没有体现出一致性的倾向。

关于个人信息界定模式,从世界范围以及中国现行规范看,并不存在统一的做法。仅从数量上看,采用单纯定义的方式居多,但这并不意味着单纯定义的方式就是最合理的界定模式。作为个人信息保护法中的核心概念,如何精确地界定其范围在很大程度上决定了个人信息保护立法的科学性和合理性。笔者认为,按照一般的法律概念界定方式,除了在概念中指出其本质特征外,对于过于抽象的法律概念,还应当列明典型类型,以实现法律概念的确定性并为社会民众提供相应的行为预期。从欧美先进国家的立法和司法经验看,由于个人信息概念的抽象性,如果仅仅规定其内涵,无疑会增加法律适用上的不确定性。因此,有必要将明确符合个人信息概念的信息,如身份证号码、社会安全码、基因信息等可以唯一识别到某个自然人的信息予以列举,从而增强个人信息概念的具体性和可适用性。

(二)个人信息外延界定

无论是美国的个人可识别信息还是欧盟的个人数据,二者在内涵上呈现出一致化的倾向,都以“可识别性”为其核心构成要件。尽管如此,要想清晰地界定个人信息的范围仍是一件十分复杂和困难的事情。个人信息保护法是调整发生在信息主体和信息管理者之间的,在个人信息收集、处理和利用等活动过程中,因保护信息主体的权益而发生的社会关系的法律规范的总称[5]66。随着信息处理技术日臻成熟,越来越多的行为,包括收集、存储、打标签、用户画像、数据分享、跨境传输等行

为均被确立为数据处理行为。因此,个人信息范围的界定将在根本上决定什么样的处理行为要受到法律的约束。对此,世界上大多数国家和地区均对个人信息的范围进行了一定程度的界定。

在美国,由于不存在统一的个人信息保护法,个人信息的定义也仅仅体现在 COPPA、《视频隐私保护法(VPPA)》《金融服务现代化法案》,以及各州的数据泄露通知法之中。在范围上,COPPA 规定,个人信息是指关于某个人的个人可识别性信息。《视频隐私保护法(VPPA)》采取同义反复的方式界定个人信息,即个人可识别信息是可以识别某个人的信息。《金融服务现代化法案(Gramm – Leach – Bliley Act)》采取反向排除法,即将非公共性(Nonpublic)的信息认定为个人信息,而各州的数据泄露通知法均采取了列举方式来界定个人信息范围,并且各州的规定并不相同。此外,比较有代表性的是,美国《消费者隐私权法案(草案)》规定,个人数据指处于受管辖实体控制之下的,通过合法方式无法被公众获取的,而且链接到或切实可由受管辖实体链接到特定个人的,或链接到个人相关的或常规使用的设备的任何数据。应当指出的是,美国不仅没有形成统一的个人信息范围,并且由于各州多采用列举的方式,导致个人信息局限于"已识别(identified)"信息,对于"可识别(identifiable)"信息则无法提供有效的法律保护[6]。

欧盟作为世界个人数据保护法的领路人,早在 1995 年就通过了《数据保护指令(95/46/EC)》(以下简称"95 指令"),第 2 条明确规定,个人数据是指任何与已识别或可识别的自然人("数据主体")相关的信息④。而作为欧盟委员会的内部咨询机构,第 29 条工作组(以下简称"WP29")就曾在 2007 年针对欧盟各成员国认定个人数据标准不一致的问题作出了《关于个人数据概念的意见》。该意见对个人数据的四大要素,即自然人、相关性、可识别性、任何信息进行了细致的分析,并列举了相关例子。最后,WP29 认为,在个人信息范围的界定上,应当结合指令的目的,即保护自然人在个人数据处理过程中的隐私权,来予以综合认定,同时考虑"所有可能的合理方法"[7]。2012 年之后,为构建欧盟一体化的数据保护规则并推动数字化单一市场的形成,欧盟委员会发起了数据保护改革,并于 2016 年通过了 GDPR,在个人数据定义上依然沿袭了 95 指令的规定,同时也界定了可识别的标准,即通过参照诸如姓名、身份证号码、定位数据、网络标识符等一项标识,或者是通过参照一个或多个针对该自然人的诸如身体、生理、基因、心理、经济、文化或社会身份因素来识别个人⑤。此外,GDPR 还在第 9 条规定了特殊种类个人数据的保护,对于揭示种族或民族出身、政治观点、宗教或哲学信仰、工会成员的个人数据,以及以唯一识别自然人为目的的基因数据、生物特征数据,还有健康、自然人的性生活或性取向数据的处理应当被禁止⑥。在 GDPR 的序言部分,欧盟再次指出,应当考虑所有合理可能的因素来认定个人数据,并明确区分了假名化和匿名化数据,前者可以通过其他额外信息来识别一个人,因此可以成为个人数据;而匿名化数据不再识别一个人,故不受 GDPR 约束⑦。从其法律规定可以看出,欧盟地区的个人数据范围相对抽象和宽泛,这也与其最大限度地保护人格利益的立法目的相契合。

英国曾于 1998 年通过了 DPA,与欧盟 WP29 不同的是,英国信息委员会办公室(以下简称

④Data Protection Directive, Article 2(a).

⑤General Data Protection Regulation. Article 4.

⑥General Data Protection Regulation. Article 9.

⑦General Data Protection Regulation. Recital 26.

"ICO")在判断什么是个人数据的方式上采取了相反的方式,即首先判断何种处理行为下的数据是法案目的下的数据,然后在此基础上判断这些数据是否是个人数据。在个人数据的认定上,ICO 同样采取了四要素分析法,并采取宽泛的保护方式。在 2017 年 9 月 14 日发布的《数据保护法令(Data Protection Bill)》中,个人数据被界定为任何与已识别或可识别的在世的人相关的信息⑧。虽然在表达上与欧盟略有不同,但二者的实质内涵一致。

德国在个人数据保护上具有其自身特色,其调整对象包含了公务部门和非公务部门的个人数据处理。在个人数据保护立法上,德国联邦宪法法院通过判例创设的"信息自决权"为其个人数据保护提供了理论基础,并于 2003 年制定了 BDSG,而在欧盟通过 GDPR 之后,又在 2017 年 6 月 30 日通过了新的数据保护法案,成为欧盟地区第一个将 GDPR 本土化的成员国[8]。在个人数据的界定上,法案完全采纳了 GDPR 的规定⑨,以保持与欧盟在未来执法上的一致性。

新加坡于 2013 年 1 月起正式施行《个人数据保护法(PDPA)》,其中在概念解释部分规定,个人信息是指,无论真实与否,能通过该信息识别或通过该信息与其他企业已经或能够获取的信息结合后识别出个人的信息⑩。

澳大利亚于 1988 年颁布实施《隐私法》,在 2016 年的修订版本中,个人信息被界定为关于一个已识别或可合理识别的人的信息或观点,无论该信息或观点是否真实,也无论该信息或观点是否以有形形式记录⑪。

俄罗斯于 2006 年通过了《个人资料法》,该法案在 2014 年经过了较大调整,在 2017 年最新修改的《个人资料法》中个人资料(亦即个人信息)是指能直接或间接地识别出或可识别出自然人(个人资料主体)的任何信息⑫。

中国台湾地区采用个人资料的概念,并在"个人资料保护有关规定"第 2 条采取列举和归纳的方式界定了个人资料的内涵⑬,同样采取可识别说。中国香港地区的《个人资料(私隐)条例》第 2 条对个人资料的内涵进行了列举,即直接或间接与一名在世的个人有关的或从该资料直接或间接地确定有关的个人的身份是切实可行的,该数据的存在形式令予以查阅及处理均是切实可行的。中国澳门地区由于历史缘故,在个人资料保护立法上更接近欧盟地区,第 4 条将个人资料界定为与某个身份已确定或身份可确定的自然人("数据当事人")有关的任何信息,包括声音和影像,不管其性质如何以及是否拥有载体。

总结其他国家(地区)个人信息保护法规发现,在个人信息概念的界定上,几乎所有国家(地区)都采取了可识别说,并且从全球影响力看,欧盟地区采取的宽范围保护模式得到了其他国家(地区)的效仿,尤其是亚太地区。作为英美法系国家的代表,英国虽然至今都没有承认一般的隐私权[9],

---

⑧Data Protection Bill. Part 1 — Preliminary. Terms relating to the processing of personal data.

⑨Federal Data Protection Act. Section 46.

⑩Personal Data Protection Act 2012. Part I Preliminary. Section 2 Interpretation.

⑪Privacy Act 1988. Part II—Interpretation. Division 1—General definitions. Section 6.

⑫Федеральный закон от 27 июля 2006 г. N 152 – ФЗ "О персональных данных". Статья 3.

⑬中国台湾"个人资料保护有关规定"第 2 条:个人资料,指自然人之姓名、出生年月日、国民身分证统一编号、护照号码、特征、指纹、婚姻、家庭、教育、职业、病历、医疗、基因、性生活、健康检查、犯罪前科、联络方式、财务情况、社会活动及其他得以直接或间接方式识别的该个人之资料。

但是其直接跳过隐私权,而采用更有包容性和时代性的个人数据保护法模式也使其立法受到关注。美国作为最早提出隐私权的国家,一直致力于隐私权的扩张,以此来保护各种人格利益,但是此种做法也具有明显的局限性,尤其是在保护范围上过于狭窄,并且呈现出分散化特点。在网络与现实生活联系日益紧密,个人信息的商业价值日益凸显的今天,此种固守传统隐私路径的做法似乎难以为个人提供充分的保障。

在中国,《民法总则》第 111 条的规定为个人信息保护奠定了基础。而关于个人信息范围的界定,目前基本上形成了以可识别说为核心的概念。其中,最具代表性的莫过于 2016 年通过并在 2017 年 6 月起实施的《网络安全法》。该法第 76 条对个人信息进行了界定,即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。该定义也与目前国际社会,尤其是欧盟地区的定义基本一致。

除此之外,中国其他部门规章以及国家标准等也对个人信息进行了界定。比如工信部 2013 年发布的《电信和互联网用户个人信息保护规定》第 4 条就使用了“用户个人信息”的概念,将其限定在电信业务经营者和互联网信息服务提供者在提供服务的过程中所收集的能够识别用户的信息,并采取列举的方式,将用户使用服务的时间、地点等信息纳入其中。应当指出的是,用户使用的时间和地点并非严格意义上的个人信息,只有在与其他信息结合能够识别用户身份时才能被认定为个人信息,该规定直接将其与其他可识别信息并列的做法也不符合个人信息的基本理论。此后,工信部于 2014 年又发布《电信和互联网服务用户个人信息保护定义和分类》,将用户个人信息界定为电信业务经营者和互联网信息服务提供者在提供服务过程中收集的能够单独或者与其他信息结合识别用户和涉及用户个人隐私的信息。该定义也反映出中国政府部门对个人信息与隐私在认识上的混淆;但是,该标准还指出,用户个人信息经脱敏处理后不纳入本标准规定的电信和互联网服务用户个人信息范围。此外,该标准还根据属性和类型特征,将电信和互联网用户个人信息分为用户身份和鉴权信息、用户数据和内容信息、用户服务相关信息三大类,相较于其他国家(地区)的个人数据与特殊类型个人数据二分法,此种做法相对复杂,也缺乏一定的逻辑性。

国家质检总局与国家标准委员会于 2013 年发布的中国首个个人信息保护国家标准《信息安全技术公共及商用服务信息系统个人信息保护指南》在个人信息的界定上就更加科学,即可为信息系统所处理、与特定自然人相关、能够单独或通过与其他信息结合识别该特定自然人的计算机数据。从其定义看出,该定义基本上遵循了 WP29 的观点,并且还采取了分类保护的模式,将个人信息区分为个人敏感信息和个人一般信息,这种界定也被其他政府部门所采纳⑭。国家质检总局、国家标准委员会《健康信息学推动个人健康信息跨国流动的数据保护指南》则采用了个人数据的概念,在内容上也采取直接定义方式,即任何涉及已标识或可标识自然人的信息,没有列举典型的个人数据种类。

除政府部门外,行业组织对个人信息保护以及概念界定也形成了一些自律规范。如中国科学

⑭中国人民银行《中国金融移动支付检测规范第 8 部分:个人信息保护》:个人信息(personal information),指信息系统所处理、与特定自然人相关、能够单独或通过与其他信息结合识别该特定自然人的计算机数据。

技术法学会与北京大学互联网法律中心联合发布的《互联网企业个人信息保护测评标准》,其中对个人信息的界定指能够切实可行地单独或通过与其他信息结合识别特定用户身份的信息或信息集合,如姓名、出生日期、身份证件号码、住址、电话号码、账号、密码等。该定义也强调识别的高度可行性。此外,测评标准也明确指出,不适用于经不可逆的匿名化或去身份化处理,使信息或信息集合无法合理识别特定用户身份的信息。

当前世界上大多数国家(地区)都采取了可识别说,但是在具体范围和界定方式上还存在一定的差异。总体而言,欧盟地区在个人数据界定方面较为先进,并在可识别性基础上出台相关意见,增加了认定标准,有利于法律适用的统一,值得中国借鉴。

二、个人信息概念界定之反思

在个人信息保护成为全球立法重点的今天,欧盟个人数据保护法以其逻辑性和体系性成为他国效仿的对象,中国亦追随此脚步,制定了大量有关个人信息的保护规范,并在《民法总则》中明确了个人信息受法律保护。对此,王利明教授指出,未来在民法典人格权编中还应当确立个人信息权[10]。此外,加快个人信息保护立法已经成为全社会的共识⑮。一项法律的独立性最核心的体现就是其调整对象,因此,个人信息范围的界定就成为重中之重,美国和欧盟之所以在个人信息跨境传输问题上摩擦重重,就在于双方在个人信息保护范围上不同,以至于有学者提出 PII 2.0 概念来调和双方间的矛盾[11]。在中国制定个人信息保护法之前,反思当前的个人信息界定,对于提高立法的科学性、可适用性和预见性具有重要意义。

(一)个人信息界定的场景性和动态性

正如欧盟法院在谷歌被遗忘权案中关于地域范围的解释所显示的,一个宽泛的解释将有助于防止个人被免除指令的保护⑯,这也是欧盟《基本权利宪章》第 7 条隐私权和第 8 条个人数据受保护权的应有之义。在概念界定上,欧盟放弃了以私密性为核心的隐私权界定,采取了看似更加客观的可识别标准,并且在形式上囊括了电子化和非电子化的信息,在内容上也包括了客观的描述性信息和主观性评价。更重要的是,欧盟对于可识别性标准的界定使得几乎所有的信息都有可能成为个人数据,而由于 GDPR 的高水平保护和严格的惩罚机制,企业的合规成本也将大大提高。隐私的界定具有主观性并受到地区传统文化的影响,但仔细分析就会发现,个人信息的界定也并非那么客观。

可识别性是个人信息最重要的要件。通常讲,“已识别”意味着在特定的人群中,某个人可与该群组内的其他人区别开来,而“可识别”则是指虽然某个人现在还未被识别,但有可能(Possible)做到这一点。在认定标准上,95 指令、WP29、《关于个人数据概念的意见》、GDPR 都一致表明了间接识别的重要性,即应当考虑“所有合理可能的方法(All the Means Likely Reasonably)”,只要在采取了合理方法能够再度识别个人的,那么其同样构成个人数据。相对于完全的列举式定义,这种界定看

---

⑮中国信息法学开创者之一的齐爱民教授团队的核心研究成果《中华人民共和国个人信息保护法草案(建议稿)》于 2017 年两会期间由全国人大财经委副主任、全国人大代表吴晓灵联名 30 多名全国人大代表作为一项议案提交两会,是迄今为止唯一一部提交两会的个人信息保护法建议稿。

⑯Google Spain v AEPD and Mario Costeja González.

似囊括了所有可能受到保护的个人信息,实则是将个人信息置于动态化和场景化的危险之中。主要体现在两个方面:其一是具有间接识别性能力的个人以何种范围为基准? 也就是说,应当以具备何种知识水平的人作为社会一般大众来判定某项信息是否具有间接识别性。由于与当事人社会关系的不同,相较于一般社会大众,与当事人有密切关系的近亲属、同事、朋友等更容易识别出该当事人。比如,对于"前 NBA 火箭队主力华人中锋"信息,无需特定的专业技能,一般人凭常识或简单检索即可识别该个人。但是,在其他情形下,如美国在线(AOL)公开搜索日志事件⑰,在该事件中,公司采取了一定的匿名化手段隐去了用户账号等信息,但是还是有专业研究人员和电脑爱好者通过技术手段识别出了特定用户。例如,第 4417749 号是佐治亚州的一个寡妇,另一个用户似乎在策划一场谋杀,此时,该信息对于一般人或技术水平较低的企业而言,其无法识别,就不构成个人信息,但是对于高水平技术人员或谷歌、微软等大型互联网企业而言,这又构成了个人信息。同一个客体在不同的主体面前具有截然相反的法律属性不仅不符合法律概念的客观性要求,同时也将造成法律适用的困难。其二是能够与间接识别性资料相结合或比对的资料,是否应当是一般人无需经过特别调查或支付庞大费用就容易获取的资料? 对此,根据 GDPR 序言(26),要确定所使用的方法是否合理可能,需要考虑所有的客观因素,诸如识别的时间长短和成本,数据处理时可用的技术以及未来的技术发展⑱。在 WP29 发布的意见中,工作组认为,实施身份识别行为的成本是一个需要考虑的因素,但不是唯一因素。控制者意欲达到的目的、数据处理的方式、控制者预期的获益、对个人的风险利益,以及组织性机能障碍(如违反保密义务)的风险和技术性失败都应当被考虑在内[7]。工作组还指出,这种可识别性的测试也是一个动态的过程,并且还应当考虑数据处理时的技术状况以及在数据处理期间技术发展的可能性。对于当前采取了所有方法都无法识别的数据,如果数据意图被存储 1 个月,身份识别在其生命期间可能不被期望。但如果该数据意图被存储 10 年,控制者就应当考虑在数据生命周期第 9 年可能出现的识别可能性,而这种可能性的出现就会使该数据成为个人数据。

由此可见,个人信息的认定并非像物一样稳定,而是随着拥有数据的主体、使用的场景、数据保存的期限、技术的发展而变化,这就决定了个人信息界定的场景性和动态性。而由于个人信息界定的动态性和场景依赖,导致现行法律的规定缺乏一定的可操作性。原因在于,在企业采取了匿名化处理并满足法律排除适用的条件后,很有可能会随着技术的进步或企业数据类型的增多而再度识别出用户。按照欧盟立法机构的设想,这将导致企业时而受到约束、时而却不受约束的奇怪现象,正如有学者指出的,这也不利于企业实施匿名化等隐私增强技术[12]。

(二)匿名化与个人信息

匿名化(Anonymisation)本身是指一种隐私增强技术,但是由于其后果是造成该数据无法再度识别到特定个人,不会损害自然人的人格利益,因而被许多国家的个人数据保护法排除在适用范围之外,中国《网络安全法》亦有规定⑲。但是计算机科学的发展已经表明,在许多情形下,非个人可识

---

⑰参见美国在线(AOL)搜索日志事件(https://en. wikipedia. org/wiki/AOL_search_data_leak)。类似事件还可参阅 Netflix Prize 事件(https://en. wikipedia. org/wiki/Netflix_Prize)。

⑱General Data Protection Regulation. Recital 26.

⑲《中华人民共和国网络安全法》第 42 条。

别信息可以被关联至个人,并且识别性数据可以被再识别。PII 和 non - PII 已经不是不可改变的范围,并且在被视为非个人可识别的信息和个人可识别信息之间存在着在某一个时刻可以相互转化的风险[13]。中国也有学者清晰地指出,匿名的状态是相对的,只在特定的场景中有效,原则上并不存在绝对的匿名化[14]。

在大数据时代,个人信息的商业价值日益凸显并体现在生活中的方方面面。利用个人信息,私人可以用于电子商务交易,政府可以更高效地实现公共政策的制定,企业可以更有针对性地开发和销售产品,医疗机构可以结合病人医疗记录、家庭背景、生活饮食习惯发现病因并做好预防工作。信息的自由流动对于全球电子商务和数字经济的发展具有重要意义。为了更充分地利用数据,企业往往采取匿名、差分隐私等手段来规避法律约束。但正如上述个人信息商业模式所展示的,其最终目的还是为了锁定至个人,并提供更符合个人需求的产品或服务。因此,识别是企业数据处理的终极目的。

与匿名化技术相关的是假名化(Pseudonymisation),二者的区别在于假名化后的数据还可以通过特定的算法或函数与原数据相关联。在性质上,有学者认为它们在欧盟法律体系中扮演着免除法律适用和法定义务的安全港角色,但从当前的技术发展趋势看是十分具有挑战性的[15]。而 WP29 特别指出,假名化降低了数据库与数据主体原始身份的关联性,就此而言,它是一个有用的安全措施,但并不是一种匿名化的方法[16],采取假名化的数据仍然是个人数据。而究竟在何种程度上,匿名化的数据才不受个人数据保护法的约束,WP29 指出,需要评估技术的坚固性(Robustness),并给出了三个具体的判定标准:(1)是否仍有可能挑出个人;(2)是否仍有可能将一个人与记录相连接;(3)有关个人的信息是否可被推测出。需要指出的是,使用此种判定并非一劳永逸,即使被认为是安全的技术,也要定期评估残余风险(Residual Risks),如果评估结果牵涉"一项有关数据主体身份识别不可接受的风险",那么即使采取了匿名化技术,对此类数据的处理仍应受到 GDPR 的约束。

可见,匿名化数据的再识别风险使得个人信息的界定更加不确定。事实上,WP29 此前就已经指出欧盟的数据保护框架是以"风险为基础的方法(Risk - Based Approach)",企业的法定义务随着处理数据的类型和对数据主体的隐私风险而变化[17]。这种方式导致匿名数据与个人数据之间的界限并非像通常所描述的那样清晰,而是随着数据环境的变化而相互转化[18]。这不仅导致法律适用上的困难,还使得企业无法有效地实施数据挖掘以及共享行为,阻碍信息的自由流动。

三、中国个人信息的立法选择

欧盟个人数据保护立法所采取的宽泛保护模式将造成法律概念界定和法律适用上的混乱,无论是从法律的确定性还是安定性角度而言,都需要通过其他制度来弥补此种不足。中国法治的现代化深受大陆法系国家影响,这一点在个人信息保护领域亦是如此。《民法总则》的出台明确了个人信息应作为一种法定利益受到保护,未来的人格权编中将确立个人信息权,为各项具体个人信息权提供权源。此外,由于个人信息保护还涉及信息自由流动、跨境传输、执法监管、侵权救济等内容,不仅超出了人格权编所能规定的容量,也由于其需要调整个人利益与社会公共利益间平衡的复杂性,制定单独的个人信息保护法已经成为世界各国的立法趋势,中国未来亦应如此。在这种宏观背景下,结合上述对个人信息界定的立法比较和反思,笔者提出了如下解决路径。

(一)加强个人信息权顶层设计

个人信息在最初是一项法益,无论在美国还是欧盟,起初都是通过隐私权进行保护。美国通过 Whalen v. Roe 案首次肯定宪法上信息隐私权,95 指令在第 1 条也规定,尤其要保护数据主体在个人数据处理中的隐私权⑳。但是,美国至今仍通过隐私权的方式保护个人信息,而欧盟却在 GDPR 中摒弃了隐私权的表述,直接使用了"个人数据受保护的权利(Right to the Protection of Personal Data)"㉑,以区分隐私权。中国在尚未确立个人信息权时,在司法实践中也是通过隐私权来保护个人信息的㉒;但是这种保护存在一定的瑕疵,最明显的就是已经公开或者涉及公共利益的个人信息被直接排除在保护范围之外。采用更为宽泛和积极的个人信息概念不仅是国际社会的主流趋势,也是克服中国隐私保护局限的理性选择。

个人信息范围的抽象性和不确定性不仅无法为企业提供明确的行为预期,也会对自然人的人格利益造成重要影响。在很多时候,自然人根本不知道其个人信息被收集之后会被如何处理,只有在其发现人格利益受损时才会去寻求救济,而且由于信息获取和传递的渠道越来越多,导致在诉讼中个人很难举证证明对方侵犯了其个人信息权;因此,确立各项具体个人信息权利对于预防个人信息滥用、减少人格利益损害就成为一种必然。在此方面,欧盟已经在 GDPR 中专章规定了数据主体的权利,包括了访问权、更正权、删除权(被遗忘权)、限制处理权、数据可携权、反对权,赋予自然人在个人数据的收集、处理、变更、删除、转移等各个环节的控制力,并且相较于其他国家,其首创的数据可携权在某种程度上类似于所有权,除了允许用户下载其提交的个人数据之外,还允许自然人将其个人数据从一个网络服务者处无障碍地转移至另一个服务者处[19]。

就中国目前而言,虽然很多法律法规都规定用户有权更正错误信息和删除违法收集的信息,但是相较于欧盟、英国等个人信息保护水平较高的地区,中国的个人信息权利体系还有很大的完善空间。在中国未来民法典人格权编和个人信息保护法中,应当以增强控制力为目的,加强个人信息权利的顶层设计,明确规定决定权、访问权、更正权、封锁权、删除权甚至是报酬请求权等权利,以克服因个人信息的动态性所导致的人格利益危险。

(二)事前知情同意

在个人信息保护法中,知情同意原则是信息管理者在收集个人信息时,应当对信息主体就有关个人信息被收集、处理和利用的情况进行充分告知,并征得信息主体明确同意的原则[20]。该原则的意义在于在个人信息的收集环节就充分告知其使用目的和相关风险,在征得个人的实质同意后实施个人信息的处理。该原则在中国现行个人信息法律法规中均有类似规定。比如《全国人大关于加强网络信息保护的决定》第 2 条、工信部《电信和互联网用户个人信息保护决定》第 9 条均要求网络服务提供者在收集个人信息前应当充分告知收集和使用规则,并征得用户同意。在个人信息保护法基本原则体系中,知情同意是实现自然人信息自决的重要方式,只有权利人真实地控制个人信息的使用,才能实现保护个人尊严的立法目的。

---

⑳Data Protection Directive. Article 1.

㉑General Data Protection Regulation. Article 1.

㉒典型案例包括:上诉人北京百度网讯科技有限公司与被上诉人朱某隐私权纠纷,江苏省南京市中级人民法院宁民终字〔2014〕5028 号;王某诉张某某名誉权、隐私权纠纷案,北京市朝阳区人民法院朝民初字〔2008〕10930 号等。

知情同意原则在实践中最重要的体现就是企业隐私政策。在《网络安全法》颁布之前,隐私政策并未受到企业的重视,存在注重形式而非实质上的同意、用户控制力保护不足等问题。此前顺丰和菜鸟物流数据之争事件就反映出快递公司在数据共享方面没有履行充分的告知义务,导致消费者根本不知道自己的数据被分享给了谁[21]。2017 年 9 月,由中央网信办等四部委启动隐私条款专项工作,针对微信、淘宝等 10 款网络产品和服务的隐私条款进行评审。此后,高德地图、微信、新浪微博等相继修改隐私政策,并通过弹窗等方式履行告知义务,这对于带动行业个人信息整体保护水平,形成社会示范效应具有重要意义。中国未来应当参考借鉴欧盟地区的做法,采用清晰易懂的语言增强隐私政策的可读性,以获得网络用户的实质同意。

(三)事中风险评估

在大数据时代,数据的自由流动是数字经济发展的前提和动力,政府和企业也在积极探索大数据交易平台并制定了相应的交易规则。根据国家标准委员会发布的《信息安全技术数据交易服务安全要求》征求意见稿,在数据交易中应当遵循个人信息保护原则,在禁止交易的部分,该标准明确指出,涉及个人信息的数据禁止交易,除非获得了全部个人数据主体的同意明示,或者进行了必要的去标识化处理。在现实生活中,个人信息被收集之后,企业很有可能会采取匿名化、跨境传输、数据共享等处理行为,即使是一方对其所收集的个人信息的特定要素采取删除、加密等措施,也有可能因为接受方所拥有的多种数据类型或技术手段而变得可识别。笔者认为,在企业实施数据交易等行为时,应当引入风险评估机制,定期对自身的数据匿名化程度、数据交易方再度识别个人的风险进行评估,以确保个人信息保护法的排除适用。

一个与之相随的问题就是,此种风险在何种程度上被视为不受个人信息保护法约束。在本质上,个人信息从完全无法识别到可识别再到已识别,反应的是识别的风险;但由于个人信息认定的场景性,从完全无法识别到可识别之间并不存在十分清晰的界线,因而导致不同地区在识别风险大小问题上存在争议。从欧盟地区发布的意见中可以看到,其对匿名化持较为严格的态度,并试图将再识别风险降低到零。英国 ICO 认为,匿名化数据并非完全没有风险,而是其应当能够降低身份识别的风险直至很微小[22]。对此,有学者指出,在匿名化没有造成任何莫须有的损害或危难的情况下,无需为匿名化处理行为正名[15]。笔者赞同此种观点,个人信息的保护不应当以牺牲信息自由流动为代价,如果企业为了匿名化付出了合理成本并将风险降低到足够程度,那么就应当在法律上视为实现了完全的匿名化。

(四)事后个案认定

个人信息的动态性和场景性决定了在不同的情况下,对一项信息的处理和使用并非一成不变。对此,中国有学者指出,个人信息的定义是动态且高度依赖于具体场景的,无法做静态的类型化界定[15]。笔者赞同此种观点,现有的法律规定虽然采取了定义加列举的方式,但是此种方式并不代表所列举的内容就绝对属于个人信息。比如在同名同姓的情况下,姓名作为一种常见的识别方式在该种情境下就不能被视为个人信息。因此,在司法实践中认定是否存在个人信息侵权时,应当坚持个案认定原则,结合案件的具体情景,在考虑所有再识别的手段后仍无法识别时,方能排除个人信息保护法的适用。

四、结论

在大数据时代,个人信息保护法成为捍卫人格利益的最后一道防线,以欧盟为代表的个人数据保护立法模式已经成为个人信息保护的主流做法,可识别性作为认定标准能够有效克服隐私界定的主观性,但个人信息界定的动态性和场景性使企业和个人在法律适用面前无所适从,个人无法知悉和控制有关其信息的处理,企业也因再识别的风险而面临高额的合规成本。从个人信息的顶层设计入手,强化事前的同意、事中的风险评估,以及事后的个案认定,能够有效弥补个人信息概念本身存在的不足。在中国未来民法典人格权编和个人信息保护法立法过程中,应当以上述原则为指导,构建科学的个人信息保护体系,以维护公民的人格利益。

参考文献:

[1]张山. 全球信息数据量逐年猛增 IDC 产业迎来发展新机遇[N]. 中国证券报,2015 - 08 - 05.

[2]齐爱民. 中华人民共和国个人信息保护法示范法草案学者建议稿[J]. 河北法学,2005(6):2 - 5.

[3]人民网. 全球 90 个国家和地区制定个人信息保护法律[EB/OL]. (2017 - 08 - 10)[2017 - 09 - 14]. http://world.people.com.cn/n1/2017/0810/c1002 - 29463433.html.

[4]张新宝. 从隐私到个人信息,利益再衡量的理论与制度安排[J]. 中国法学,2015(3):38 - 59.

[5]齐爱民. 信息法原论[M]. 武汉:武汉大学出版社,2010.

[6]SCHWARTZ P M,SOLOVE D J. Reconciling rersonal information in the United States and European Union[J]. California Law Review,2014,102:891.

[7]ARTICLE 29 WORKING PARTY. Opinion 4/2007 on the concept of personal data[EB/OL]. (2007 - 06 - 20)[2017 - 09 - 14]. http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf.

[8]BERND M W,LECHNER F. New German Federal Data Protection Act[EB/OL]. (2017 - 08 - 07)[2017 - 09 - 14]. https://www.paulhastings.com/publications - items/details? id = 4dddec69 - 2334 - 6428 - 811c - ff00004cbded.

[9]王泽鉴. 人格权的具体化及其保护范围·隐私权篇(上)[J]. 比较法研究,2008(6):1 - 21.

[10]王利明. 论我国《民法总则》的颁行与民法典人格权编的设立[J]. 政治与法律,2017(8):2 - 11.

[11]PAUL M,SCHWARTZ D J S. Reconciling personal information in the United States and European Union[J]. California Law Review,2014,102:905.

[12]WALDEN I. Anonymising personal data[J]. International Journal of Law and Information Technology,2002(10):225.

[13]SCHWARTZ P M,SOLOVE D J. The PII problem: Privacy and a new concept of personally identifiable information[J]. New York University Law Review,2011(86):1814.

[14]范为. 大数据时代个人信息定义的再审视[J]. 通信保密,2016(10):70 - 80.

[15]ESAYAS S Y. The role of anonymisation and pseudonymisation under the EU data privacy rules: Beyond the 'all or nothing' approach[J]. European Journal of Law and Technology,2015(6):1 - 5.

[16]ARICLE 29 DATA PROTECTION WORKIING PARTY. Opinion 05/2014 on Anonymisation Techniques[EB/OL]. (2014 - 04 - 10)[2017 - 09 - 16]. http://ec.europa.eu/justice/data - protection/article - 29/documentation/opinion - recommendation/files/2014/wp216_en.pdf.

[17]ARICLE 29 DATA PROTECTION WORKIING PARTY. Statement on the role of a risk - based approach in data protection legalframeworks[EB/OL]. (2014 - 05 - 30)[2017 - 09 - 12]. http://ec.europa.eu/justice/data - protection/article - 29/

documentation/opinion – recommendation/files/2014/wp218_en. pdf.

[18]STALLABOURDILLON S,KNIGHT A. Anonymous Data v. personal data – false debate: An EU perspective on anonymization, pseudonymization and personal Data[J]. Wisconsin International Law Journal,2016(34):321.

[19]张哲. 探微与启示:欧盟个人数据保护法上的数据可携权研究[J]. 广西政法管理干部学院学报,2016(6):43 – 48.

[20]张才琴,齐爱民,李仪. 大数据时代个人信息开发利用法律制度研究[M]. 北京:法律出版社,2015:42.

[21]王林. 菜鸟顺丰掐架敲响警钟个人信息怎么保护?[N]. 中国青年报,2017 – 06 – 06(09).

[22]UK ICO. Anonymisation: managing data protection risk code of practice[EB/OL]. (2014 – 04 – 10)[2017 – 09 – 16]. https://ico. org. uk/media/1061/anonymisation – code. pdf.

Identification and reidentification: The definition of personal information and the legislative choice

QI Aimin, ZHANG Zhe

(*Law School, Chongqing University, Chongqing 400045, P. R. China*)

Abstract: In big data era, data has become a key element in social production and economic development, and there is enormous potential value in personal information in improving government decision – making, enterprise' s precision marketing, innovation of social management and so on. Globally, the legislation of personal information protection law represented by the European Union has become the mainstream practice in the international community. As for the definition of personal information, the international community has basically formed an approach based on identifiability. But the dynamicity and contextual nature of the definition of personal information not only leads to the difficulty of judicial application, but also makes the enterprise not be able to deal with the problem of anonymization appropriately. In order to make up for the inherent defects of personal information definition, it is necessary to take full advantage of foreign legislation, strengthen the top – down design of personal information rights and construct a prior consent, risk evaluation and case – based mechanism, which constitute the proper meaning of promoting the rationality of civil code and personal information protection law in the future.

Key words: personal information; re – identification; anonymization; right to personal information

(责任编辑　胡志平)