

## 第十二章 尊重隐私和保守秘密：履行契约

### 一、对于调查对象的政策

522. 在制定调查对象政策过程中,最重要的一点或许是统计机构通过尊重调查对象来赢得公众对它的信任,而不仅仅把尊重调查对象当作实现统计目标的手段。牢记这一点很重要:即使在法律要求调查对象对一项或多项数据收集工作作出回应的时候,公众参与统计机构的调查在很大程度上依然是一种自愿行为。即使调查工作并非自愿性的,统计机构依然有义务以合乎职业道德的态度对待调查对象,也就是说,尽量减轻他们在时间上的负担,尊重他们的隐私,信守调查对象提供信息时统计机构做出的保密承诺。

523. 尊重隐私意味着承认正是个人“拥有”关于自己的信息。隐私的概念指由个人来决定提供什么信息,什么时候提供,以及向谁提供。要求因统计目的而提供私人信息的法律得以通过的前提是公众对该信息的需求压倒一切,同时这些法律含有条款,避免泄漏数据的可识别性。这一概念同样适用于公司等法律实体。当然,公有公司同意定期公布销售额、利润等信息以便参与规范的股票市场。私营公司或股权不外让的公司应公开的信息则要少得多。

524. 统计机构在获得隐私数据时恰当的应对措施应当是保证数据的机密性。调查对象提供隐私数据时,履行调查对象和统计机构之间的这一契约是一项承诺。本章将讨论这一契约关系的要点以及从调查对象手中获得数据时保护数据的手段。

525. 至关重要是要制定政策,以此促使属意调查的调查对象形成合作的心理。下面是这些政策的要点,我们对每个要点进行解释。

#### 1 与调查对象关系的原则

526. 调查对象有两大类:一类是企业,另一类是个人或住户。有的原则对两者都适用,而有些则仅适用于企业。(1)项到(5)项同时适用于两类,而(6)项和(7)项只适用于企业。

##### (1) 数据收集的对于调查对象一定要明确和具有意义

帮助调查对象了解收集数据的原因常是一项困难的任务,而这一原则对于小型企业尤其困难。收集数据的总体框架使得统计机构无法频繁地进行普查。实际上,考虑到效率,抽取的样本量最好不要超过能够迅速和准确地得出汇总结果所必要的范围。但是,一个小型企业特别关注的是该企业营业地区的情况以及定义范围很窄的活动,在大多数统计工作中,这两方面的信息都不具备充分的代表性,除非采用普查或大规模的全面调查。对于两类调查对象,解释目的时都必须做到尽量清楚和有说服力。

##### (2) 统计机构必须让外界认识到是根据最严格的保密原则掌控着所有的个体记录,避免它们落入政府内或政府外的其他方面手里

统计机构必须作出这一保证,使调查对象感到放心<sup>108</sup>。这一保证包含两点:必须让调查对象了解,统计机构掌握的数据不会落入怀有不良企图的人手中<sup>109</sup>;法律认可由统计机构掌握的保密数据不会同该国的政治当局、管理机构或参与政策制订工作的公务员机构共享。一些国家还特别表示,司法部门不得下令交出提交给统计机构的数据。

##### (3) 统计机构必须被认为乐于配合调查对象,例如作出进一步的解释或者同意用合理的方式替代传统的调查问卷

在企业统计中,统计机构不仅应该知道企业是如何记帐的,还必须愿意不怕麻烦地提取

<sup>108</sup> 有记录的案例表明,企业明确宣布它们愿意由国家统计局收集甚至是常规的信息,条件是所提供的信息免于未经授权单位查看。

<sup>109</sup> 当然,这种保证是有限度的。但是公众应当得到保证,穿透统计机构设置的防御是极端困难的。

出所需要的数据，尽量减少企业花费的力量。例如，一份股东报告已经包含了数据收集人员所寻求的全部变量或大部分变量，所以应该可以用来替代一份要填写的统计问卷。

#### **(4) 统计机构的专业水平、客观性、免于政治干预的地位必须确立、接受和不断宣扬**

在走向调查对象的时候，统计机构提出的要求不得超过必需的范围，而且在此之前就必须已经决定获得信息的最少负担的方式。这一要求的后盾是机构的声誉，特别是它所用来获得结果的方法能够维护其自主性和免于政治干预。

#### **(5) 统计机构应当被认为是对调查负担问题有深思熟虑和深切关注；也就是说，它应当承诺要找到办法来简化书面工作的负担**

总的来说，应当要认识到，统计机构为了获取信息而走向调查对象应是最后一招，即所有的渠道都经过考虑但都不能充分解决。统计机构对于调查负担问题要有一项负责的政策，并务必使数据收集人员彻底理解政策中的最后一招的含义。统计机构尤其要表明以下几点：

- 统计机构一直密切监控其信息收集工作；
- 统计机构已知道它的工作必然会打搅和侵犯隐私，其结果是慎重平衡好尊重隐私和促进公开研讨所需最低限度的信息之间的关系；
- 收集到的信息不会被滥用；
- 它能够评估造成的调查负担的程度，并且可以通过实施相关政策来减少负担。<sup>110</sup>

527. 此外还有两条原则，只适用于企业群体：

#### **(6) 收集信息的方法必须遵从企业做记录的方法**

成功的信息收集工作能够把企业所持有的记录转化成数据编制者所需的标准记录。但是，这项工作应当由统计机构而不是企业来完成。如果数据收集人员了解企业是如何做记录的，就可以判定信息的转换在多大程度上没有扭曲。总的来说，统计机构最好拥有众多透彻了解企业会计和记帐方式的员工。最后，应当通过努力使企业相信，向统计机构提供机密信息是安全的。

#### **(7) 收集的信息应当采用企业日常工作中使用的术语**

区别两种语言十分重要，一种是统计人员内部交流使用的语言，另一种是同调查对象沟通的语言。“活动种类单位”和“净利润”等表达方式对统计机构以外的人来说几乎没有什么意义。用于获取信息的语言必须是调查对象所熟悉的语言，无论这个调查对象是一家住户，还是一家企业。<sup>111</sup>

528. 然而，统计机构不是唯一收集信息的机构。实际上，政府出于不同的原因，通过多种途径从公众那里收集详细的信息。例如，为了征税或者提供补贴而收集详细的资产负债表信息；为了执行与健康和安全有关的法律和规定而收集详细的商品信息<sup>112</sup>。统计机构应当是减少卷面工作最积极的追求者，只要由法律或者因为内容上的重叠给出机会，统计机构就应使用其他部门收集到的信息。它应当能够向公众表明，它在同其他政府部门沟通，如何利用法律允许的手段来防止重复收集，特别是对那些基础十分不完备而无法填报的企业更是如此。

<sup>110</sup> 某些机构目前每年都计算对企业所造成的总体负担。根据这项计算，统计机构向政府提交一份正式报告，说明上一年所减轻的负担总量。

<sup>111</sup> 澳大利亚统计局作出了全面的努力，试图用“行业使用的”语言全面改写企业调查问卷。这项工作的投资相当可观。但是即使它不能使回答率有显著提高，至少有助于减少回答误差。

<sup>112</sup> 经合发组织的两个成员国在不同时间进行的两项统计表明，统计数据收集过程产生的书面工作量占政府全部书面工作量的比例稍稍超过 5%。

## 2 强制和自愿回答

529. 国家与国家之间的法律的本质有所不同。在某些国家，配合统计收集工作完全是一种义务。如果企业或住户成员等调查对象不按照要求的形式按时提供信息，它们就触犯了法律。在其他一些国家，法律上要求提供某些种类的信息，而其他要求则基于自愿。最后，在某些情况下，法律对此问题含糊不清。如果出现这种情况，统计机构可能害怕对调查对象提出太多的要求，因为如果受到挑战，法律可能会裁决调查对象没有义务提供任何信息（人口普查内容除外），而裁决结果一旦公布，将对回答率带来负面影响。

530. 大多数国家可能属于第二种情况，法律只认可一些有限的义务调查。无论法律基础如何，所有统计机构都会发现，最重要的目标是获得调查对象的合作，特别是小型企业和住户的合作。<sup>113</sup>

531. 如果有选择的余地，统计机构可以采取一种适中的姿态，即经济调查是强制性的，而其他调查则是自愿性的。无论法律如何规定，牢记这一点很重要：没有合作的态度，有多少强制措施也不能缓解回答率这一难题。

## 3 激励和保证

532. 在许多国家越来越严重的一个问题是数据收集过程中的回答率低下。已经尝试过的可能的解决办法之一是对调查对象采用激励措施。然而，激励措施并不总能解决问题。在某些国家，使用经济刺激是非法的。在另外一些国家，预算过紧，无法采用大量激励措施。而且，如果对调查对象（如住户成员）采用激励措施，他们需要花费的时间和补偿的性质之间必然存在某种联系。某些调查可能很琐细。例如，在态度调查和连续调查中，住户成员所付出的代价就是住户的私人事务定期要受到干扰。在另外一些情况下，尤其是对于消费者收支调查，住户要花费的精力较大<sup>114</sup>。最后，付酬激励措施会引起调查对象定下期望值，没有激励措施的调查就难以开展，或者调查对象会不断地要求提高报酬。以下几点应当考虑：

- 选择的激励措施必须具有显著效果而且不过头，使调查对象不会对它不屑一顾，决定统计机构预算的人也不会把它当成一种浪费。现场调查组织在对此提供意见方面扮演重要角色。
- 无论选择何种激励措施，最好让接受调查的住户成员了解采用或不采用激励措施的原因，了解该统计机构的形象，当然还要让他们意识到，提供信息帮助研究与公众利益相关的问题是一种公民义务。<sup>115</sup>

533. 尽管激励措施用得最多的对象是住户成员，在对小型农业业主和小型非正规企业的调查中也起作用。这一点不一定适用于大型企业，因为提供相关信息本身对它们来说就是一种有力的激励措施。例如，如果典型案例中编用的企业信息十分全面，可以被用于省级细分项，在全省运营而非地区运营的企业就可以将它们在该省的成就和同一层次的竞争者进行比较。所以，以提供进行比较的专门表格作为一种激励措施——只要数据相当细化和足够新颖——就足以使得企业界与统计机构之间有一种建设性的合作关系。<sup>116</sup>

534. 然而，在所有统计中，统计机构都必须坚决而明确地承诺，提供出去的信息不会泄露调查对象的身份。对任一规模的企业而言，坚决的保证加上认知，即被要求提供的信息是国家进行有序管理所必需的，就足以让企业放心。换言之，在企业调查中采用激励措施通常是不必要的（而且出于预算方面的原因，激励措施也不符合统计机构的最佳利益）。

<sup>113</sup> 总的来说，大型企业会响应政府的要求，并且不会要求政府详细地解释为什么要获得某类财务和经济信息。然而，有些跨国公司在总部所在国以外的地方对提供信息上采取不合作的态度。

<sup>114</sup> 在许多情况下，收支调查同固定资产和金融资产调查结合进行，从住户处收集资产负债表、现在和以往交易等详细数据。这不仅要求住户详细记录日常交易情况，而且调查员本身也要花大量时间帮助他们完成调查问卷。

<sup>115</sup> 在履行公民义务方面，提供私人信息按出席陪审团一样对待。

<sup>116</sup> 法国的国家统计及经济研究局（INSEE）成功地实行一种激励企业的政策，那就是提供按客户要求订制的列表，包括企业和同行间具体比较表。

## 4 分享行政信息

535. 与统计机构分享详细的行政信息有许多先例，最值得注意的例子是海关。然而，税收机构也收集大量信息，也像统计学家一样注重信息的完整性和保密性。分享信息，特别是分享税收机构的信息，是一件棘手的事情。如果要分享信息，税收机构必须得到保证：统计机构至少会像他们一样，非常小心地处理个体信息。以下是统计机构需要牢记的要点：

- 无论统计学家和收税员之间分享的是什么信息，信息只能朝一个方向流动——从税务部门流向统计学家；
- 为了让税收机构乐意分享信息并采取合作的态度，统计学家需要为税务部门提供一些服务，前提是不损害至关重要的保密问题；
- 税务部门收集的信息是减少小型企业的回答负担的最有效的资源；<sup>117</sup>
- 税务部门感兴趣的是按照他们的经济活动分组对纳税人提交的表格进行分类，分类方法不一定和统计机构使用的方法相同；
- 税务纪录的行业分类方法如果与用于统计的分类方法相一致，将带来重大的战略性收益；
- 公众对统计学家将税务纪录进行分类的看法，可能会削弱统计学家做出的为提交给他们的信息保密的承诺。

536. 统计机构面临的挑战是寻找出一种最好的折衷方案，既要满足这些条件，又要在任何情况下不损害统计的保密性要求。所有符合税务部门要求的方法同样适用于通过行政手段生产信息的其他部门——偶尔会有所不同。

537. 除了行政信息的提供者所要求的各种保证——无论是协议形式还是口头承诺——以外，统计机构还应当表明，敏感信息的物理存放会受到非常严密的保护。在许多统计机构，进入这些地点要受到监控，外界人员如果想进入存放税务文件的楼层需要获得特别授权。

## 二、数据保护

### 1 物理和电子安全措施

538. 目前，统计机构中有两个领域需要加以保护。第一是物理领域，即文件、计算机纪录、缩微胶片、照片等物体的实际存放地点。在这方面，如果记录被保存在一个以上的建筑物中，统计机构不仅会面临较高的保护记录的成本，而且还需要说服公众相信，它们这样做同样是为了个人记录的安全。第二是虚拟领域，即电子领域，统计机构内部通信系统对其进行跟踪，并与个体信息的储存连接在一起。

539. 这两部分都需要加以保护，以免遭人恶意侵入或者未经授权的侵入。如前所述，在许多统计机构，访问信息受到严密监控（例如，员工必须佩带证件），如果保存非常敏感的信息，还要采取其他安全措施。

540. 电子领域也必须受到保护，既有密码准入，还有他安全措施，但是这一点更难做到，尤其是技术不断地要求升级防御系统。公众媒体大量报道黑客成功渗入到最安全的通信系统并窃取最机密的信息，有时甚至涉及到国家安全文件。由于公众担心计算机安全的程度超过对传统文件安全的担心，而且这种趋势似乎在加剧，一些统计机构选择彻底把内部通信系统与外界隔离，包括同其在地方上的办事处隔离。

541. 在这些统计机构，数据储存和与外界通讯之间没有直接联结。而且，无论什么保密资料，如果要从地方办事处传输到中央总部都必须有最好的加密系统的保护。现在还无法预测这种方法是否可以在技术迅猛发展的情况下提供足够的保护。统计机构的系统分析人员中必须有一位被指派专门负责数据安全，负责加密、消除表明身份的特性、防止恶意篡改资料等事项。此外，还要专门指派人员负责场地和数据存储的物理安全。

<sup>117</sup> 那些只有少数人口交纳所得税的国家除外，这种情况下考虑统计目的也没有意义。

## 2 处罚

542. 某些官方规定必须说明对损害安全的处罚措施，以此作为一种威慑，并向公众表明统计机构严肃对待所拥有数据的保密性和完整性的问题。最通行的方法是通过法律本身。理想的情况是由法律条文明确规定各种违法行为。按照问题的严重程度的顺序，举例说明这些违法行为：

- **疏忽。**调查者把一个信封遗忘在公共汽车上，内装有可识别个体数据的已完成的调查问卷，第三方可能发现这个信封，而且新闻界可能把这当成是公共部门管理松懈的一个例子；
- **不当行为。**某一统计专业的专家在公共场合评论住户收支调查中某特定家庭所表明收入（造成的后果和上一条相同）；
- **恶意行为。**某位希望让统计机构的管理层感到难堪的员工把总理的人口普查表格送给新闻界；
- **为牟取私利使用保密信息。**某一员工把一公司的保密信息提供给它的一个竞争者。

543. 这一切当然仅仅是例子。法律可能会更加全面地规定什么行为构成安全上的犯罪行为。然而，这些例子表明，有多种不同的违法行为，应该列举出来，并受到切实惩罚。

## 三、保密和泄露

### 1 原则

544. 以上各节已经讨论了对有关收进的信息的防范措施，如何使它不落入他人手中，以及如果出现损害信息安全行为时采取什么样的惩戒措施。本节将讨论与调查对象达成的契约应当如何得到履行。我们将考虑一些方法，以避免我们发布了汇总数据，但却泄露了某一个人、企业或机构身份的信息<sup>118</sup>。公布的信息越详细，犯下这种错误的风险就越大。例如，通常根据地域、调查对象的人口特征以及收入来公布一项家庭支出调查的结果。即使调查获得数以万计的答复，统计数据中属于极端范围的调查对象（如富有的名流）的身份可能很容易地识别出来。

545. 标准的解决办法是开展一些初步研究，确定哪些单元会识别出调查对象。在接下来的研究中，确定出最低汇总程度的数据，从而遮蔽<sup>119</sup>调查对象的身份。通常，这种研究包括两个层次。首先，采用一种完全是实践的、基本的方法，经过反复试验，以最低的代价来排除这些有问题的案例。对能够负担得起的统计机构而言，还有一种更富有理论的方法来找出有问题的单元，并说明所提出的解决方案在遮蔽信息方面代价最小。由于很少有统计机构具备此类研究的能力，另一种解决办法是委托相关的学术机构来设计最低限度数据遮蔽模型。当然，研究结果也可以应用在其他地方，而且鼓励小型统计机构同大型统计中心保持沟通——尤其在这方面。

### 2 支配和残差泄露

546. 一些与保密有关的问题一直是过去 50 年间反复讨论的对象，随着计算机技术广为传播和用户群因此而扩大，这些问题变得更加重要。存储、交叉列表和公布统计数据的能力的提高不经意间增加了意外泄露的机会。此外，用户也提高了运用数据的能力，提高了在数据本身之外发现更多信息的能力。<sup>120</sup>

<sup>118</sup> 一个有趣的问题是，这些是唯一需要保护的信息类型吗？比方说，如果我们公布心脏病专家的平均收入，或者不同种族的犯罪率，这会违反契约吗？有关到底什么信息应当受到保护的问题现在还未成为慎重审视的对象。

<sup>119</sup> 使用“遮蔽”一词是要说明，不是要排除基本信息，而是应当匿名。解决办法是把单元合并起来。

<sup>120</sup> 假设某一统计机构根据需求执行一项制表业务，其内容是从一项复杂的调查中得出机器可读格式的表格。总体原则是任一用户所得到的任何东西均对所有用户敞开。我们再假设有  $n$  个用户，每个用户希望得到不同的交叉列表。结果是统计机构将不得不分析第  $n$  个要求对可能的残差泄露数量的影响，假设前提是

547. 对于每一份列表，统计机构都必须跟踪弄清每个单元包括有多少个调查对象，以及每个单元中有多少调查对象可以以残差的形式计算出来。必须采取措施来遮蔽低于门槛的所有数字（下限通常是三个）。在调查对象的数量等于或者超过门槛的单元中，有一、两个调查对象占所显示的变量数值的90%以上，也应该采取措施遮蔽它们的活动。

548. 对那些有大量出版计划的统计机构而言，有理由为跟踪支配和残差案例<sup>121</sup>的相关工作设立一个专门单位来开展对于保密问题的研究。这一单位通常是负责使用数理统计部门的分支机构（至少与之紧密相关）。<sup>122</sup>

### 3 保密和住户统计

549. 一个人或者住户可能和其他人或住户很相似，但是索尼、皇家荷兰壳牌石油公司集团、国际商用机械公司、德比尔斯公司、俄罗斯天然气工业股份有限公司和通用汽车公司却与小型企业大不一样。鉴于这个原因，社会统计和经济统计有若干不同之处，我们必须予以考虑。在社会统计中，只有在住户资产、收入及支出的调查中才涉及到支配性这一概念。此外，在经济统计领域中，任何一个调查样本一般都包含某一类单位，因为它们在变量中占很大一部分，但家庭和个人只要是非零概率即可抽选。最后，除了专门调查（如住户财务资产负债表）以外，可以在住户调查的结果中引入随机扰动，从而遮蔽个体案例，但不会影响对任何分布动差的估计值。然而企业调查中很少出现这种情况。

550. 最近，跟踪调查引出了一些与保密研究有关的有趣问题。这些调查长期跟踪一个抽样单位（一般为一个人或家庭），从而可以观察其特征的演变。这类调查在比较发达的统计机构中越来越普遍。此类调查在时间上的特征提高了样本被识别的可能性。这些调查要求开展数据分析和描述，但是最低限度必需的遮蔽措施可能会导致这些分析和描述毫无作用。显然，对住户统计的保密研究与企业统计研究有很大不同，所以应当作为研究议程中的一个单独课题。

### 4 获得同意的公开

551. 就公众形象而言，统计机构使自己处于报表负担监护人的地位，不断表明其愿望，即在可能的情况下要尽量减少文书工作量，这符合统计机构的最大利益。相应的，统计机构应当主动寻找机会，利用其他各方面法定权力来获得信息，而不应重复政府其他机构已经进行过的调查。然而，避免重复调查可能意味着分享信息，这包括两种方式：统计机构从政府部门收集的数据中获取信息，或者反过来。如果有选择的话，调查对象出于众多原因通常希望回答统计机构的问题，其中最重要的原因是统计机构在谨慎处理和注重细节方面享有声誉。如果可能分享信息，调查对象自己必须同意（最好以书面形式）让第三方查阅他们的报表，统计机构还必须让调查对象了解为什么会出现分享信息的情况。还应当告诉调查对象，如果他们断然拒绝，结果可能是他们不得不分两次提供相同的信息。

552. 可能还有一些其他情况，需要调查对象以书面形式同意公开信息。在高度工业化国家，如果严格按照字面意思解释防泄露规定，那么行业分析表（如投入产出表）可能会丧失其价值。在接受牺牲统计汇编的分析价值之前，统计机构可能希望企业同意承担经营活动被公开的风险。可以询问这些企业是否同意某种形式的公开，而这种公开会导致有缺陷的秘密防护。当然，大规模地采用这种做法可能会损害统计机构的声誉。

### 5 被迫公开

553. 在某些情况下，除了公开某一特定企业的活动之外别无其他可行的选择。这种情况大

---

第 n 个用户了解先前 n-1 个要求的内容。

<sup>121</sup> 要全面了解跟踪和遮蔽制度，请参见 Gordon Sande 的《维护商业统计秘密的自动化单元遮蔽》，发表于《统计数据库管理第二次国际研讨会论文集》，Los Altos, 加利福尼亚，1983 年 9 月 27-29 日，Roy Hammond 和 John L. McCarthy 编辑，(Lawrence Berkeley 研究室，1983 年)。

<sup>122</sup> 该部门的工作一般包括加密研究。

多发生在公用事业、国家垄断企业和行业，它们由某一公司控制，在小国家中常常会遇到这种情况。如果这种情况广为人知，而且实际上在统计机构成立之前就存在了，如何处理这一情况的专门条款通常会被写进法律。在年度统计中，这一情况问题不大，因为国有企业和国家大型垄断企业通常要非常详细地报告它们的经营活动和财务状况，而且这些报告是公开的。

## 6 被动保密

554. 完全避免对秘密的侵害是不太可能的，最为明显的例子是国际贸易统计。在国际贸易统计中，收集到的数据是记录在案的交易，而不是进行交易的个人或者企业。由于这些统计形成了对相关时期交易情况的普查，非常详细地公布这些统计数据，并且包括若干项属性（交易的商品、目的地、运输手段、进出港口等），所以认为能够做到完全保密并且不会严重影响所提供信息的时效性和实用性是不现实的。惯用的方法是选择一种主动和被动相结合的保密方法。在所有只知道一家进口商或者出口商的情况下<sup>123</sup>，或者在具有支配地位的情况下，可以设计出并事先实施遮蔽措施。在所有其他情况下，只有在企业或者个人主动向统计机构提出抱怨时才采取遮蔽措施。如果统计机构遭到指责未能及时对问题采取行动的话，就应该让调查对象认识到，与统计机构在设计和实施调查的工作中采用推理方法发现信息泄露的情况不同，国际贸易统计使用对外行政工作汇编的记录，因而统计机构不负有直接的责任。所以，只有在接到通知的时候才能采取行动。

## 7 研究工作

555. 在特定情况下，根据统计机构自己的命令，统计机构可公布抽中企业的名称、地址、产业活动（用产业编码表示）和规模（用雇员数组或某种其他公认的变量，还有规模组）。一般来说，这种例外处理用于研究工作的目的。

556. 完全正当的应用研究现在越来越多，即不仅使用统计汇总数据，而且还使用个体数据，以便成功地开展研究工作。例如，通过分析已公布的统计数据，对于消费收入、支出、住户储蓄和税收之间复杂互动关系的研究常常受制于很大的汇总偏误。在个体跟踪记录的基础上对正式的假说进行检验可能更加易于进行。没有什么单位凭借自身的力量就能进行此类研究工作。所以在这种情况下，如果统计机构向研究人员提供他们需要的信息，允许他们提出假说和进行必要的研究，那么就能实现共同的利益。

557. 现在，统计机构无法回避在某些情况下要提供微观数据的问题。然而，建设一个含有个体信息而又可以进入的数据库，其中定然能识别身份的概率应该是极低或者是不存在的。当然，所有的记录都应当匿名，而且只要原始分布动差得以维护，就可以有引入随机扰动的基础<sup>124</sup>。即使在这些扰动以后，仍然要指定一名官员，负责监管进入数据库的有关事宜以及由此派生出来的工作。这意味着统计机构中有一名官员，始终了解谁出于什么目的在使用数据库，在寻找什么，到现在取得了那些结果。一些用户要依据特定的详细数据才能开展工作，统计机构必须协调好同这些用户之间的关系。这样，广大公众就会相信，这些详细数据有助于促进社会科学的发展。

558. 很难想象公众可以进入的数据库能够提供企业信息，因为每个产业一般都有一个产业单位或者公司，其规模很大，足以支配数据结果；其对数据的影响程度高，足以被识别。然而，小型企业的情况可能不同，而且如果有一些问题只与小型企业有关，建设一个面向公众的数据库则是合理的。请注意，尽管以住户为内容的公用数据库获得成功的例子不胜枚举，但以企业为内容的却寥寥无几。

559. 研究界对数据的特殊需求不容忽视，尤其在那些需求被认为是要服务于公众利益的时

<sup>123</sup> 这种方法运作起来是不对称的。大多数国家的进口商远比出口商多。后者可以轻易地提前辨别出来，因此可提前采取措施遮蔽可识别的信息。但是所有企业都可以看作是进口商，所以很难作出先验性的判断。

<sup>124</sup> 如何提供不可辨别但含有和原始数据集相同的信息是一项正在积极研究的课题。参见 D.B. Rubin 的《超过 18 岁以后的多重推算》，载于《美国统计学会杂志》第 91 卷，473—489 页。

候。其中许多需求合理地提出要接触个体记录，或者接触列表中的具体单元，而这些单元中调查对象的数量少于门槛的规定。由于此类研究不是为了个人目的，也不是要获取商业情报来损害既定利益，所以统计机构很难做到不同意。相反，统计机构同样难以不顾向调查对象作出的承诺，允许某些用户有例外。一个统计机构很少能有效地平衡好这两种压力的关系。下面是一些统计机构采用的一种行之有效的解决办法。

560. 一种在某些国家获得有限成功的办法是正式聘用要求获得微观数据的学术研究人员，尽管他们的报酬只是象征性的。这种办法保证研究人员同意遵守统计机构的规则，知道不这样做会受到官方惩罚。保证这种正式聘用成功的条件如下：

- 研究人员拥有合适的证书；
- 统计机构认为研究计划涉及的内容符合公共利益，或者是为了促进社会科学的进步，而且没有其他可替代的数据来源；
- 研究人员同意在聘用期内维护该统计机构的协议和立场。

561. 在提出聘用之前，统计机构应当确保该研究人员符合上述三个条件，或许还要咨询其法律顾问。同样重要的是不要利用这一做法来绕过保密规则。很清楚，滥用这种做法，或者为了微不足道的项目采用这种做法，无疑都会损害统计机构的信誉。谁拥有该研究人员的研究成果的知识产权，这个问题不属本《手册》的讨论范围。

562. 另一种方法曾在美国人口普查局试验过，包括设立一个“无菌室”供研究人员工作。研究人员不具有雇员的身份，但是完全了解统计机构的保密政策，并在违反保密政策时接受法律制裁。

## 结论

统计机构承诺对清楚显示个体调查对象的运营、财产、态度或其他任何特征的信息进行保护。这样做的原则很清楚，但是有一些具体应用方法需要不断研究，以保证不违反保密性问题。理论研究由于可以分担并且反复使用，因而很有价值，但是成本很高，只能在少数统计机构中进行。但是即使在此类研究开展不了的情况下，也应当始终做好更注重实践研究的工作，既要遮蔽显性数据，又不要损害统计汇总数据的完整性。这些活动的运作应当公开，使之成为让公众放心的一种手段。

绝对保密规则的确也有例外。如果发生这种情况，统计机构应该勇于面对公众，并且清楚进行特殊处理的条件。住户记录和企业记录不要以同样的方法处理，因为前者更加容易遮蔽，而后者经常出现如果不损害发表的数据的质量就无法加以遮蔽的情况。在这些情况下(如自然垄断企业或支配性企业)，要么统计法包含有一条公开信息的专门条款，要么统计机构必须在征得可能被识别的调查对象明确表示同意的前提下，才能采取行动。