



统计委员会

第三十五届会议

2004年3月2日至5日

临时议程* 项目6(e)

未按领域划分的活动：统计数据和元数据的列报

经济合作与发展组织关于统计数据和元数据的列报问题的 报告

秘书长的说明

秘书长根据统计委员会第三十四届会议的要求,** 谨转递经济合作与发展组织关于统计数据和元数据的列报问题的报告。报告简要介绍了关于编写一份手册的各种提议。这份手册拟载列关于国家机构和国际组织在各种传播媒介中提供的统计数据和元数据的列报准则和建议采用的最佳做法。之所以需要编写一份统一现有标准并酌情提出新建议的手册,是因为需要进一步改进数据质量(特别是可解释性和一致性),并尽可能减轻向国际组织报告数据和元数据的负担。

报告第七节载有请委员会提供指导意见的具体问题。

* E/CN.3/2004/1。

** 见《经济及社会理事会正式记录,2003年,补编第4号》(E/2003/24),第一章,A节。



经济合作与发展组织关于统计数据和元数据的列报问题的 报告

目录

	段次	页次
一. 导言	1-3	3
二. 需更明确地订立数据和元数据列报标准的背景情况	4-11	3
A. 改进数据的可解释性和一致性	5-8	3
B. 尽可能减轻报告负担	9-11	4
三. 国家和国际两级现有列报标准的相关性	12-14	5
四. 列报准则手册的可能范围	15-19	5
五. 与有关国际倡议之间的联系	20-27	7
A. 统计数据和元数据交换项目	21-22	7
B. 在国际一级拟订共同变量清单	23-25	8
C. 国民账户世界交换项目	26-27	8
六. 制订指导方针手册的拟议时限	28	9
七. 寻求委员会予以指导的问题	29	9
附件		
拟议中的数据和元数据列报手册可能包括的内容		13

一. 引言

1. 统计工作协调委员会于 2003 年 9 月 8 日至 10 日在日内瓦举行的 2003 年会议上，讨论了经济合作与发展组织（经合组织）编写的一份文件，其中概述了关于编制一份手册的提议草稿。这份手册拟简要说明关于国家机构和国际组织在各传播媒介中提供统计数据和元数据的列报准则和建议采用的最佳做法。

2. 为便于统计工作协调委员会在 2003 年 9 月讨论与数据和元数据的列报有关的问题，经合组织就数据和元数据的列报标准开展了初步工作。之所以需要在经合组织范围内拟定此类标准，最初是因为需要减轻国家提供者向经合组织报告数据的负担，并改进数据质量。经合组织从这一初步工作得出结论认为，任何列报手册的编写都需要目前正在这一领域开展工作的其他国际组织和若干国家机构的积极参与。统计工作协调委员会欢迎经合组织提议拟定国际数据和元数据的列报手册，并请经合组织就这项提议编写一份文件，供统计委员会第三十五届会议讨论，文件应重点说明是否需要拟定一套综合的国际列报准则，并载述有关这份手册可列有哪些内容的一些初步想法。

3. 本报告：

(a) 提供背景材料，其中强调需要在国家和国际两级制订有关数据和元数据列报的准则；

(b) 简要说明任何列报手册可能涉及的范围；

(c) 列出与目前正在国际一级开展的其他有关活动的联系；

(d) 列出编写准则手册的时间表草稿，以便为委员会 2005 年会议编出初稿；

(e) 说明请委员会予以评论和指导的若干领域。

二. 需更明确地订立数据和元数据列报标准的背景情况

4. 国家和国际两级均存在两个主要的迫切问题，表明需提出一整套关于数据和元数据的列报标准。它们是：需要改进数据质量，以及在向国际组织提供数据和元数据方面尽可能减轻报告数据的负担。

A. 改进数据的可解释性和一致性

5. 第一个紧迫问题，是需要改进向国家和国际两级的用户提供的统计数字的质量，特别是改进可解释性和一致性（数据集内部、不同的数据集之间、先后提出的数据之间、以及不同国家之间数据的一致性）。这些都是质量的重要方面，一些国家机构和国际组织（如欧洲共同体统计局(欧统局)、国际货币基金组织(货币基金组织)、经合组织、加拿大统计局等等）拟订的每一种质量框架中已经以

这种或那种形式包含了这些重要方面。但是，此类框架除了说明需要作出改进之外，很少详细解释在列报数据和元数据时，将如何执行这些质量方面。

6. 可解释性反映的是用户可在何种程度上不费力地理解并适当使用和分析数据。数据所涉概念的定义、目标总体、各种变量和所用术语、以及说明数据可能具有的限制的资料等是否适当，大体上决定了可解释性的程度。如元数据的列报与各不同用户的需求以及结构合理、随时可用的数据的用途相符合，对可解释性也有帮助。

7. 关于一致性，用户在比较同一机构在不同时间编纂的统计数字以及不同国家的机构和不同国际组织编纂的统计数字方面往往遇到三个主要问题：

(a) 因使用不同的变量定义、单位和分类而在概念上产生的差异；

(b) 因不同国家在数据收集和处理做法方面的差别而在操作上产生的差异；

(c) 数据列报的不同做法。此类差异包括：列报的类别（绝对数字、指数、增长率）、数据形式（未经处理的数据、按季节调整后的数据，趋势-周期数据）；订正做法；抽样误差的列报；基准年；因可随时查阅适当的统计元数据而在方法上具有透明度。

8. 拟议编写的手册将侧重于第三个问题，并力求在一个框架内列出主要的列报做法，同时提出供国际组织和国家机构在以不同形式提供产出时采用的建议、准则和最佳做法草案。

B. 尽可能减轻报告负担

9. 第二个紧迫问题，是需要尽可能减轻国家机构向国际组织提供数据和元数据的负担。最近，国际论坛（如欧洲统计员 2002 年会议（经合组织/货币基金组织，2002 年）和经合组织高级别统计小组 2003 年会议（经合组织，2003 年））在讨论中，就扼要说明了国家来源与不同的国际组织之间在传送数据及其相关的元数据方面采用数据分享模型会有哪些惠益。这种模型设想由国际组织从国家机构网站所列的数据中选取所需的共同数据。

10. 过去五年里，由于新技术、特别是网基技术的发展，已经有可能在技术上采用数据分享模型。采用这一做法的前提条件是，不仅要解决信息技术方面的若干技术问题，而且需要国家机构和国际组织就若干数据“内容”问题达成协议，包括：

(a) 确定主要统计数字所需的一套共同数据（见下文第 23-25 段）。本报告下文第五节简要说明了旨在推动国际组织协调一致地从国家来源收集数据和元数据的若干相关倡议，这些倡议也直接或间接地有助于上文所述数据分享模型的发展；

(b) 商定关键数据列报做法，这些做法应有助于查明国家机构和国际组织提供的相同数据序列，并有助于特别是国际组织提供相容一致的数据。

11. 这些领域的所需准则要由国际组织与国家机构合作加以拟定。显然只有在其数据库载存可分享的数据和元数据的国家机构的积极参与下，数据分享模型才能得以执行。数据分享意味着国际组织之间相互协调，从而使数据的提供发生根本性变化，也使国家机构通过执行不仅旨在改进数据的可解释性和一致性，而且还有助于数据的提供的数据和元数据列报准则向国际组织提供数据的作用发生根本性变化，从而最终减轻国家机构的报告负担。

三. 国家和国际两级现有列报标准的相关性

12. 拟议编写的手册拟提出的建议和准则，将借鉴国际组织与国家机构合作拟订的范围广泛的现有国际统计标准（联合国统计司，2002a 和欧统局，2003b）。这些标准主要侧重于概念问题，但也包含收集数据方面的定义问题、分类、所涉范围和最佳做法。对于数据列报问题，国际标准基本上没有谈及，或仅简单地提了一下。¹ 即使货币基金组织的特别数据提供标准（货币基金组织，2003b）也仅泛泛地提到列报标准，而且偏重于说明需提供元数据以提高可解释性，并在数据订正方面采用良好做法。欧统局的短期商业统计手册提出，需进一步统一欧盟各成员国列报指数和增长率的方式。欧统局认为，这会有助于欧统局核对由其提供的的数据是否与国家提供的的数据序列相一致（欧统局，2002 年）。由于缺少现有的国际标准，在一些情况下有必要拟定新的建议。

13. 国家和国际两级的几乎所有机构都编有文件，其中载列技术准则，说明有关数据列报、图表编排、引用等等的编列标准，供参与编制统计出版物的人员使用。这些准则基本上针对纸面出版物的编写，对数据列报问题的说明往往非常笼统。理想的情况是，关于数据和元数据列报的拟议国际手册的关键要素都载于此类文件中，并（或）与之有关联。

14. 拟议编写的手册的目标之一，就是在这一个来源里提出关于统计数据列报以及如下文第 16 段所述对数据列报产生影响的各种做法和过程的一整套准则。鉴于此类做法可能因所用的传播媒介的不同而有所差别，可能需要专门针对在线数据库以及通过网站、纸面出版物、其他电子产品和新闻稿提供数据等各种不同传播媒介，制订具体的准则。预计今后的国际统计标准将载列与这样一份国际手册相一致或与手册有关联的列报做法。

四. 列报准则手册的可能范围

15. 预计在国际一级拟订的任何列报准则都会包含若干具体建议，涉及所有数据都需具体说明的两个主要方面，即：

(a) 数据类别——绝对数字、指数、增长率。绝对数字可以是存量序列（衡量有关活动在某一时间的情况，也可以是流量序列（衡量到某一日期为止有关活动的情况）。指数、增长率和比率是绝对数字的进一步变换形式。绝对数字可选用以下两种形式之一列报：——物理单位（数目、吨）；或——以现价或固定价格表示的价值。每年或每隔更长的一段时间发布的统计数字一般都会提供绝对数字。同时也为许多短期指标提供此种数据，如有关机动车辆登记情况、建筑许可证等等的月度或季度数据等。但更为常见的是以指数或增长率的形式提供短期统计数字，因其更便于人们就一段时间以来经济现象的变化作出结论。增长率有若干不同的类别；

(b) 数据形式——未经处理的数据（原始数据或未按季节调整的数据序列）、按工作日调整的数据、按季节调整的数据、趋势周期数据。

16. 此外，还有少数主要的数据列报做法，也会对数据的可解释性产生重要影响，而且各国家机构和国际机构就其采用的不同处理办法还使数据分享模型的执行和国家数据的比较等更为复杂。这些做法包括：

- (a) 数据的订正；
- (b) 对序列中断的列报；
- (c) 抽样误差；
- (d) 在重算指数的基数时使用共同做法；
- (e) 相关的、但非完全相同的序列/变量的列报；
- (f) 引用做法；
- (g) 维持分类的做法，包括有关分类的历史记录；
- (h) 元数据的获得及其列报（包括关于元数据列报的显式模板）。

17. 预期将在今后关于这些问题的工作中提供范例，说明现有的良好做法，并确定出少数几项主要建议。这些做法已经简要列于若干地方，其中包括货币基金组织关于标准和守则遵守情况的报告（货币基金组织，2003a），后者扼要说明了各国在何种程度上遵守了国际公认的标准，包括与提供数据有关的标准。手册中还将列入对相关工作（包括相关的学术工作）的大量文献索引。还可考虑根据手册所列的准则，编写适当的培训材料，供国家一级的讲习班使用，特别是供发展中国家的讲习班使用。

18. 有一项因素使国际列报标准的拟订工作较为复杂，这就是不同术语的使用，这特别是指不同国家使用的各种形式的增长率以及同一国家就不同序列使用的各种形式的增长率。使用不一致的标记常会引起误解。与使用不一致的术语有关的问题，还更多地出现在数据的收集以及实际编写简要介绍国家的数据收集、处

理等做法的含定义的元数据文本方面。在比较国家数据时，这种不一致严重限制了许多现有的元数据的使用。为了协助克服这些问题，若干国际组织已经拟定了范围广泛的词汇，其中含有大多取自现有国际标准的主要概念和变量的定义。经合组织的《统计术语汇编》（经合组织，2002a 和 2002b）即为此类词汇的一例，以下机构也编有其他词汇：欧统局（欧统局，2003a 和联合国统计司，2002b）。而且，正如在下文第五节的进一步讨论中所指出的那样，目前正结合统计数据 and 元数据交换项目编制另一项词汇（《元数据通用词汇》）。

19. 与数据列报标准有关的最后一个主要问题是，需要国家机构和国际组织拟订适当的元数据，说明其列报做法，并使这种元数据便于统计专业知识程度不同的用户查阅和理解。理想的情况是，由不同国家中的不同组织用共同术语表述这种元数据。拟议编写的手册也将讨论主要在哪些方面存在术语问题，其中还会列入与数据列报有关的概念汇编。

五. 与有关国际倡议之间的联系

20. 目前正在国际一级进行的几项倡议将会从（国际组织和国家机构）拟订和通过一套共同的数据列报做法中直接或者间接受益。最近关于制订数据分享模型的倡议进行了下述三个互相联系的项目。如上文第 10 段所述，将数据列报标准列入拟议手册是实施该模型的关键要素。

A. 统计数据和元数据交换项目

21. 统计数据和元数据交换项目²是由七个国际组织（国际清算银行、欧洲中央银行、欧统局、货币基金组织、经合组织、联合国统计司和世界银行）组成的联合体开展的项目，该联合体致力于发展一套统计信息领域的共同商业做法，以便这些组织在现有的集体活动范围内更有效地交换和分享数据及元数据。该项目的目的是，探索哪些共同标准和正在进行的标准化活动可有助于这些组织在统计信息领域提高工作效率，避免自己的工作重复，并可能使其他组织提高工作效率和避免工作重复。统计数据和元数据交换倡议将在议程项目 6(f) 下更详细地讨论。

22. 统计数据和元数据交换倡议目前正在进行的四个项目之一是，编制一份词汇《元数据通用词汇》（经合组织/欧统局，2003 年），将其作为一项工具，帮助确保国家及国际两级编者编制的元数据就其内容及所涉方法问题范围来说均相容一致。《元数据通用词汇》的目的是，将不同国家及国际机构开发的不同的元数据模型使用的各种元数据术语都包括在内。在统计数据和元数据交换倡议项目范围内，特别注意确保《元数据通用词汇》将货币基金组织的数据公布特殊标准元数据模型的术语都包括进去，尽管该词汇还准备供其他国际组织和国家机构开发的元数据模型使用。

B. 在国际一级拟订共同变量清单

23. 国际一级目前采用的共同变量清单的最明显例子是，经合组织、欧统局、货币基金组织、世界银行和联合国统计司用来收集年度国民账户数据的调查表。调查表概列这些国际组织查明的一套共同的国民账户数据要求，其主要目的是减轻各国机构的报告负担，在国际一级提供相容一致的数据。调查表由一套十分详细的国民账户变量组成，这些变量是为了达到国际机构的要求而专门规定的，以一种共同准则和特定的编列格式（例如，以本国货币按时价/不变价格计算），列在一整套 Excel 表格中。调查表构成了下文所述国民账户世界交换项目的基础。

24. 但是，就拟订共同变量清单而言，年度国民账户调查表基本属于例外情况，尽管所采用的方式可以用于那些需要一些国际组织共同、不间断地收集数据以制订结构指标和短期指标的其他统计领域。在短期经济指标方面，经合组织和欧统局拟订了独立的短期经济指标变量清单，可作为拟订国家机构特别关心的一个关键的报告负担领域的共同变量清单的起点。这些清单包括：

(a) 经合组织制订的列入其每月公布的主要经济指标数据库的“目标”指标清单。此清单的主要目的是，使经合组织向成员国机构和其他国际组织提出的关于提供主要经济指标数据及方法方面的信息的要求有一个重点。为了确保收集一套有尽可能多的国家使用的“共同”指标，必须有这样一个重点。随着优先次序不断变化，用户关心的新问题不断出现，每隔一段时间，就对清单进行修订。没有任何一个经合组织成员国编纂清单所列的全部指标；

(b) 规定了欧盟成员国提交给欧统局的数据的所涉期间和形式的《欧洲联盟委员会短期统计条例》（欧洲联盟委员会，1998年）中所规定的变量清单。2001年又确定了一份清单，载有一套对欧洲统计系统提出的优先短期经济指标要求，即欧洲主要经济指标。随着时间的推移，将不断完善该清单。该清单还包括目标公布日期及其他质量目标。欧统局将在成员国通过欧洲指标网站进入的共用传播平台上提供的资料的基础上汇编并公布欧洲主要经济指标，将覆盖欧盟/欧元区地区以及按照欧盟标准汇编的国家指标。

25. 有必要查明类似所有有关国际组织制订的年度国民账户调查表的一套共同的短期经济指标变量。2003年6月召开的经合组织统计问题高级别小组会议上承认有必要这样做，并呼吁国际组织共同努力制订这样一份短期统计清单。该清单将包括各种所需变量，而且最好包括采用数据分享模型时这类数据的列报形式。

C. 国民账户世界交换项目

26. 2002年召开的经合组织国民账户专家会议建议进行一项实验，以检验国民账户领域数据交流模型在各国家机构和经合组织的实施情况。国民账户世界交换项目（经合组织，2002c）意在实施一种模型，在该模型中数据不是在各组织之间

传输，而是在万维网上公布，而且所采用的形式使用户只需输入国家参照数和变量参照数即可提取数据。这样做的目的是，从上文所述国家机构为了向国际组织递交年度国民账户数据而已编制的 Excel 表格开始。

27. 国民账户世界交换项目的另一个目的是，使国际组织收集的数据成为国家机构正式提供的数据。该模型的两个好处是，将向国际组织提交报告的负担降低到最低限度；最大限度地提高国际统计界的数据的质量，因为他们使用的数据是正式提供的数据，而不是专门为国际组织编制的并向它们提供的数据。

六. 制订指导方针手册的拟议时限

28. 在编撰载有一套初步的数据和元数据列报形式指导方针和建议在关键领域采用的良好做法的文件草稿方面，经合组织已经开展了一些初步工作。在今后 12 个月里，将继续进行这项工作，把统计工作协调委员会就手册拟涉及的问题提出的建议以及统计工作协调委员会成员的其他意见，连同统计委员会就本报告下文第七节概述的问题提供的指导意见并入。有关未来编写该手册的拟议时间表如下：

	重要日期
编写一份草案，供出席 2004 年 6 月 28 日和 29 日在巴黎召开的经合组织短期经济统计专家组会议的经济组织成员国初步审阅	2004 年 4 月 30 日
进一步的工作及并入短期经济统计专家组的评论	
提交给统计工作协调委员会供其在将于 2004 年 9 月 6 日至 8 日在纽约举行的下一次会议上讨论	2004 年 7 月 30 日
并入统计工作协调委员会的评论	
提交给统计委员会第三十六届会议（2005 年），以征求意见	2004 年 12 月 17 日
并入统计委员会的评论，编写手册的定稿	2005 年 6 月 30 日

七. 寻求委员会予以指导的问题

29. 寻求委员会对以下方面提出和指导：

- 编写数据和元数据列报国际手册的理由和必要性（如上文第二节概述的）；

- 拟议手册打算包括的内容，查明目前在上文或者本报告附录（概述拟议手册可能的内容）中没有述及的数据和元数据列报问题，特别是那些与年度统计和/或社会统计有关的问题（如上文第三节概述的）；
- 关于在编写手册时征求国际组织和国家机构的意见的建议（如上文第六节概述的）；
- 查明国家机构或者国际组织编写的有关列报问题的任何相关参考材料，如果委员会认为这些材料会有助于编写该手册。

注

¹ 存在例外情况，例如，《欧洲联盟委员会短期统计条例》（欧洲联盟委员会，1998年）规定了送交欧统局的数据的所涉期间、类型和形式，例如，绝对值、指数、非季节性调整、趋势——周期等等。但是，各项条例不打算详细论述列报形式问题，或者规定通过文档移交的方式向欧统局提供数据。

² 关于统计数据和元数据交换倡议更详细的情况以及在该倡议总体下目前正在进行的四个项目，可进入 www.sdmx.org 内该交换倡议网站查阅。

参考文献

欧洲联盟委员会(1998年)。关于短期统计数据的1998年5月19日第1165/98号《理事会条例》。<http://forum.europa.eu.int/irc/dsis/bmethods/info/data/new/1165-98en.pdf>, 2003年7月29日存取。

欧统局(2002年)。《短期商务统计方法:解释与指导方针》。http://forum.europa.eu.int/irc/dsis/bmethods/info/data/new/embs/MM_Eurostat%202002-0507%20vo.2.pdf, 2003年7月29日存取。

- (2003a)。欧统局概念和定义数据库。<http://forum.europa.eu.int/irc/dsis/coded/info/data/coded/en.htm>, 2003年7月29日存取。
- (2003b)。欧统局分类服务器。www.oecd.org/findDocument/0,2350,en_2649_37423_1_1_1_1_37423,00.html, 2003年7月29日存取。

Finn, Brian (2003年)。《关于数据列报标准术语的一些提议》,为2003年6月26日和27日在巴黎召开的经合组织短期经济统计专家组编写的讨论文件。<http://www.oecd.org/dataoecd/12/60/2789033.pdf>, 2003年8月7日存取。

货币基金组织(2003a)。《关于标准和守则遵守情况的报告》。www.imf.org/external/np/rosc/rosc.asp, 2003年7月29日存取。

- (2003b)。数据提供标准公告栏。<http://dsbb.imf.org/Applications/web/dsbbhome/>, 2003年8月1日存取。

经合组织(2002a)。《统计术语词汇》。<http://cs3-hq.oecd.org/scripts/stats/glossary/index.htm>, 2003年7月29日存取。

- (2002b)。《经合组织<统计术语词汇>概览》,2002年10月在巴黎召开的经合组织国民账户专家会议上提出的讨论文件。<http://www.oecd.org/dataoecd/57/34/1946320.doc>, 2003年8月7日存取。
- (2002c)。《国民账户世界交换》,2002年10月在巴黎召开的经合组织国民账户专家会议上提出的讨论文件。<http://www.oecd.org/dataoecd/10/62/1953780.doc>, 2003年8月7日存取。
- (2003年)。《经合组织统计信息新系统》,2003年6月13日在日内瓦召开的经合组织统计问题高级别小组会议上提出的讨论文件。www.oecd.org/dataoecd/27/57/2955756.pdf, 2003年7月29日存取。

经合组织/货币基金组织(2002年)。《国际组织数据和元数据收集方面新发展的进度报告》,2002年6月10日至12日在巴黎召开的欧洲统计员会议上提出

的讨论文件。<http://www.unece.org/stats/documents/ces/2002/8.e.pdf>, 2003年8月7日存取。

经合组织/欧统局（2003年）。《元数据共同词汇项目》。www.sdmx.org/General/Projects.htm, 2003年8月7日存取。

联合国统计司（2002a）。统计方法出版物。<http://unstats.un.org/unsd/progwork/>, 2003年7月29日存取。

- (2002b)。联合国共同数据库定义。http://unstats.un.org/unsd/cdb/cdb_help/cdb_quick_start.asp, 2003年7月29日存取。

附件

拟议中的数据和元数据列报手册可能包括的内容

前言

目录

1. 引言和概览
 - 1.1 拟订列报标准的必要性
 - 1.2 该手册所涉列报问题的范围
2. 与其他倡议之间的联系
 - 2.1 统计数据和元数据交换项目
 - 2.2 在国际一级拟订变量共同清单
 - 2.3 国民账户世界交换项目
3. 数据列报术语
 - 3.1 问题术语
4. 不同类型数据列报指导方针
 - 4.1 绝对数字
 - 4.2 指数
 - 4.3 增长率
5. 不同形式数据列报指导方针
 - 5.1 原始数据
 - 5.2 工作日调整
 - 5.3 季节调整数据
 - 5.4 趋势周期
6. 主要列报做法指导方针
 - 6.1 数据的订正
 - 6.2 对序列中断的列报
 - 6.3 抽样误差
 - 6.4 重算指数基数时采用共同做法
 - 6.5 相关的但非完全相同的序列和变量的列报
 - 6.6 引用做法

- 6.7 维持分类的做法，包括有关分类的历史记录
 - 6.8 元数据的获得及其列报（包括关于元数据列报的显式模板）
 - 7. 词汇
 - 8. 文献目录
-