

文件编号：EGS N300-5

日期：2005 年 12 月 15 日

标题：“政务信息资源目录体系 第 5 部分：政务信息资源标识符编码方案”

来源：“政务信息资源目录体系”工作组

内容摘要：“政务信息资源目录体系 第 5 部分：政务信息资源标识符编码方案”编制说明和征求意见稿

页数：16

电子政务标准化总体组秘书处

联系人：高栋

通信地址：北京市东城区安定门东大街 1 号

电话：(010) 84029792 传真：(010) 84029792 手机：13401035600

电子邮件：gaodong@cesi.ac.cn

《政务信息资源目录体系

第 5 部分：政务信息资源标识符编码方案》

（征求意见稿）编制说明

一、 任务来源

为落实中办发[2002]17 号和中办发[2004]34 号文件精神，推进政务信息资源目录体系与交换体系建设，受国家标准化委员会和国务院信息化工作办公室委托，由国家电子政务标准化总体组组织编写《政务信息资源目录体系》和《政务信息资源交换体系》系列标准。

《政务信息资源目录体系 第 5 部分：政务信息资源标识符编码方案》标准由中国标准化研究院、大唐电信集团、北京信息资源管理中心、中国电子技术标准化研究所、天津南大通用数据技术有限公司共同负责起草。

二、 目的和意义

随着全国电子政务建设的不断深入开展，各个政府部门纷纷建立其政务信息系统。这些系统相对孤立、分布广泛，使得大量的信息孤岛逐渐形成。为了促进部门间的信息交换和共享，首先和迫切需要的就是建立政务信息资源目录体系。

建立政务信息资源目录体系，就是要实现对政务信息资源的编目、注册、发布、查询和维护，对政务信息资源实施高效的管理。因此，为了提供对信息资源进行检索和查询的依据，有必要对每一项政务信息资源赋予一个唯一不变的标识码。

三、 国内外现状

1、 国际信息资源标识现状

国际上对信息资源的标识方案里，最著名的有美国的数字对象标识体系 (Digital Object Identifier, DOI) 和英国的信息资产登记注册库标识码 (Information Assets Register Number, 简称 IARN)。

——数字对象标识体系(DOI)

DOI 的主要目的是唯一标识网络环境下的各种信息资源实体 (Entity)，包括各种物理和数字资源。DOI 是从统一资源定位符 (Uniform Resource Locator，

URL) 发展而来, 被称为“下一代 URL”。它与 URL 的最大区别就是实现了对资源实体的永久性标识。

DOI 的编码方案(即美国标准 ANSI/NISO Z39.84-2000)规定, 一个 DOI 有两部分组成: 前缀和后缀, 中间用“/”分割。对前缀与后缀的字符长度没有任何限制, 因此理论上, DOI 编码体系的容量是无限的。

DOI 前缀由两部分组成, 一个是目录代码, 所有 DOI 的目录都是“10.”, 即所有 DOI 代码都以“10.”开头。另一个是登记机构代码, 任何想登记 DOI 的组织或单位都可以向 IDF 申请登记机构代码。登记机构代码的分配也是非常灵活的, 如一个出版商可以为其所有的信息资源只申请一个前缀, 也可以为其数字图书、音像制品各申请一个前缀。

DOI 后缀是一个在特定前缀下唯一的后缀, 由登记机构分配并确保其唯一性。后缀可以是任何字母数字码, 其编码方案完全由登记机构自己来规定。后缀可以是一个机器码, 或者是一个已有的规范码, 如 ISBN 号或 ISSN 号。

——信息资源标识体系 (IARN)

IARN 是英国皇家文书局 (HMSO) 制定的一套服务于信息资产登记注册库系统 (Information Assets Register, IAR) 的信息资源标识体系。英国皇家文书局是英国内阁办公室的下设机构, 专门负责向公众、企业和政府提供广泛的政务信息服务。信息资产登记注册库是由该机构维护的信息资源目录系统。

IARN 标识码由机构码和内部码两部分组成:

IARN 机构码是由 HMSO 分配给每个政府机构的机构代码, 如“CA000”是分配给内阁办公室的机构代码。

机构内部码由某个机构为特定信息资源分配的机构内部标识码, 在这个机构内部必须是唯一的。该代码可以是一个自动生成的机器码, 也可以是一个遵循特定标准的规范码, 比如 ISBN 号, 或者是其他已有的文件号。

2、我国信息资源标识方案发展趋势

参照国际上通行的做法, 我国在制定信息资源标识方案时, 可采用“两段式”代码结构, 即前段码加后段码的格式, 中间用“/”分割。其语法格式如下:

<信息资源标识符> = <前段码>/<后段码>

前段码是由信息资源前段码管理中心统一管理, 并由其分配给各级信息资源目录管理者, 或分配给拥有和提供信息资源的政务部门。

后段码由各级目录管理者或资源提供者自行管理,并由其唯一分配给每一个注册的信息资源。

四、 编制原则和依据

在《政务信息资源目录体系 第 5 部分：政务信息资源标识符编码方案》的编制过程中,遵循了以下原则:

(1) 以需求为导向,为政务信息资源目录体系的应用服务

在提出我国的政务信息资源标识方案时,充分考虑了我国政务信息资源目录体系建设对标识符的需求。

(2) 与国际接轨,兼顾国家信息资源标识方案

在标准的编制过程中,标准起草组对国际上的信息资源的标识方案进行了大量的深入的研究、对比和分析。最后通过对国际上比较成功的 DOI 和 IARN 两种编码方案进行参考,并通过对我国信息资源的发展趋势进行研究,确定我国政务信息资源标识符的编码方案。

(3) 符合政务信息资源目录体系,具有系统性的原则

在标准的编制过程中,标准起草组与政务信息资源目录体系的其他部分标准起草组进行充分的信息沟通、研讨与协调。保证第 5 部分与其他部分的一致性以及整个目录体系的系统性。

五、 工作过程

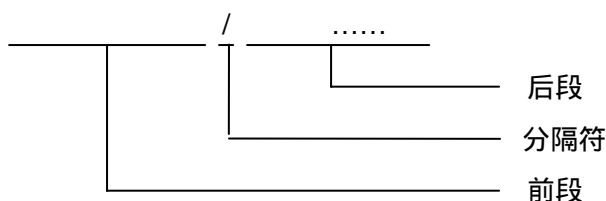
任务下达后,标准起草组立即开展了政务信息资源标识方面的研究。标准起草组认真研究了国际上 DOI、IARN 等信息资源的标识方案,认真研究了我国信息资源标识的现状和发展趋势,之后标准起草组初步提出了我国政务信息资源标识的方案。

从 2004 年底启动该标准的制定以来,课题组先后召开了 20 余次标准讨论会,其中大型封闭修改 4 次。在多次听取国家电子政务总体专家组以及标准化领域专家对标准的修订意见和建议后,经起草组认真的研究和讨论,对这些意见进行了合理化处理,最终在 2005 年 9 月初形成了标准的征求意见稿。

六、 主要技术内容

《政务信息资源目录体系 第 5 部分：政务信息资源标识符编码方案》规定了政务信息资源标识符的编码的结构、前段码和后段码的管理与分配原则。

参照国际上通行的信息资源编码方法以及我国的信息资源标识的发展趋势，政务信息资源标识符由前段码和后段码两部分组成，前段码和后段码之间用“/”隔开。前段码共 5 位，由 10 个阿拉伯数字(0 - 9)和 26 个大写英文字母(A—Z)组成。后段码长度不固定，可以采用 GB 18030 中规定的任意字符。前段码由信息资源前段码管理中心进行统一管理，并分配给目录管理者或相关的政务部门；后段码由各目录管理者或政务部门自行确定其编码方案。政务信息资源标识符的编码方案可用下图表示：



七、 与其他标准的关系

(一) 本部分与标准的其他各部分相互配合，共同形成了政务信息资源目录体系标准。其中，第 1 部分提出了目录体系的总体框架；第 2 部分提出了目录体系的技术要求；第 3 部分定义了目录体系的核心元数据；第 4 部分提出了政务信息资源的分类方案；第 6 部分规定了目录体系的管理方面的要求。第 5 部分标识符编码方案在技术上既相对独立于以上各部分，又是第 3 部分和第 6 部分的支持和补充。

(二) 本部分不同于《第 4 部分：政务信息资源分类》中的编码。第 4 部分规定的政务信息资源分类编码针对具有某些相同特征的一类政务信息资源，其目的是分类；而第 5 部分规定的政务信息资源标识符针对某一单个的政务信息资源，其目的是唯一标识。

标准起草组

2005 年 9 月



中华人民共和国国家标准

GB/T ×××××.5—××××

政务信息资源目录体系 第 5 部分：政务信息资源标识符编码方案

Government information resource catalog system

Part 5 : Rule for Government information resource Identifier

(征求意见稿)

××××-××-××发布

××××-××-××实施

中华人民共和国国家质量监督检验检疫总局

中国国家标准化管理委员会

发布

目 次

前 言.....	III
引 言.....	IV
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 标识符的结构.....	1
4.1 标识符的构成.....	1
4.2 前段码.....	2
4.3 后段码.....	2
5 标识符的管理与分配.....	2
5.1 前段码的管理与分配.....	2
5.2 后段码的管理与分配.....	3
5.3 政务信息资源标识符的生成.....	3
附录 A（资料性附录）政务信息资源标识符的生成案例.....	4
附录 B（资料性附录）前段码的分配规则.....	6

前 言

GB/T XXXXX 在总标题《政务信息资源目录体系》下，目前分为6个部分：

- 第1部分：总体框架；
- 第2部分：技术要求；
- 第3部分：核心元数据；
- 第4部分：政务信息资源分类；
- 第5部分：政务信息资源标识符编码方案；
- 第6部分：技术管理要求。

本部分为GB/T XXXXX的第5部分。

本部分由国务院信息化工作办公室提出。

本部分由国家电子政务标准化总体组归口。

本部分的主要起草单位：中国标准化研究院、大唐电信集团、北京信息资源管理中心、中国电子技术标准化研究所、天津南大通用数据技术有限公司。

本部分的主要起草人：刘碧松、魏宏、孙文峰、孙广芝、张巧英、刘守华、高栋、王国军。

引 言

随着全国电子政务建设的不断发展,各级政府部门纷纷建立其政务信息系统。这些系统相对独立、分布广泛,容易形成信息孤岛。为了促进部门间的信息交换和共享,首先和迫切需要的就是建立政务信息资源目录体系。

建立政务信息资源目录体系,就是要通过对政务信息资源的有序编目、注册、发布、查询、维护和管理,从而实现政务信息资源的共享、服务和应用。为了实现这一目的,必须对每一项政务信息资源赋予一个唯一不变的标识符。

政务信息资源目录体系

第 5 部分：政务信息资源标识符编码方案

1 范围

本标准规定了政务信息资源标识符的编码方案，以便为每一项政务信息资源分配一个唯一不变的标识符。

本标准适用于政务信息资源的编目、注册、发布、查询、维护和管理。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改单（不包括勘误的内容）或修订版均不适用于本标准，然而，鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本标准。

GB 18030-2000 信息技术 信息交换用汉字编码字符集 基本集的扩充

GB/T XXXXX.1-XXXX 政务信息资源目录体系 第1部分：总体框架

3 术语和定义

GB/T XXXXX.1中确立的以及下列术语和定义适用于本部分。

3.1

标识符 **identifier**

用于标识或命名一个数据项目并可能指出该数据确定性质的一个或一组字符。

[GB/T 14915-1994中定义]

3.2

政务信息资源标识符 **government information resource identifier**

用于唯一标识政务信息资源的一组字符。

3.3

编码 **coding**

按一定规则将一个集合的元素映射为另一个集合的元素的过程。

注1：元素可以是字符或字符串。

注2：第一个集合是编码集，第二个集合是代码元素集。

4 标识符的结构

4.1 标识符的构成

政务信息资源标识符由前段码（5 位字母数字）和后段码（不定长的字符串）两部分组成，前段码和后段码之间用字符“/”隔开，其形式如图 1 所示。

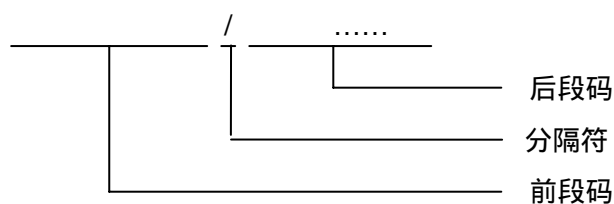


图 1 政务信息资源标识符的组成

4.2 前段码

前段码是政务信息资源标识符中的字符“/”之前的部分，用来标明为政务信息资源分配后段码的实体。该实体可以是各级政务信息资源目录管理者（以下简称“目录管理者”），也可以是拥有并提供政务信息资源的政务部门。

前段码共 5 位，由 GB 18030 中规定的 10 个阿拉伯数字（0~9）和 26 个大写罗马字符（A~Z）组成。

4.3 后段码

后段码是政务信息资源标识符中的字符“/”之后的部分，用来对同一实体内部所管理或拥有的信息资源进行唯一标识。

后段码的长度和编制规则本标准不作规定，所采用的字符可以是 GB 18030 中规定的任意字符。

5 标识符的管理与分配

5.1 前段码的管理与分配

政务信息资源标识符的前段码由信息资源前段码管理中心（以下简称“前段码管理中心”）统一管理和分配。

前段码通常分配给各级目录管理者，也可以分配给拥有政务信息资源并向目录管理者提交（或注册）这些资源或其元数据的各级政务部门。前段码管理中心负责确保前段码的唯一性，不得将一个前段码分配给多个实体。

各级目录管理者和各级政务部门应在筹建政务信息资源目录体系之初申请前段码，以便在开展目录体系建设时实现对政务信息资源的唯一标识。

目录管理者申请前段码时，首先应向上级目录管理者提出，上级目录管理者在认定符合条件后，代其向前段码管理中心申请前段码。

图 2 给出了目录管理者申请前段码的具体流程，其中阿拉伯数字表示先后顺序。

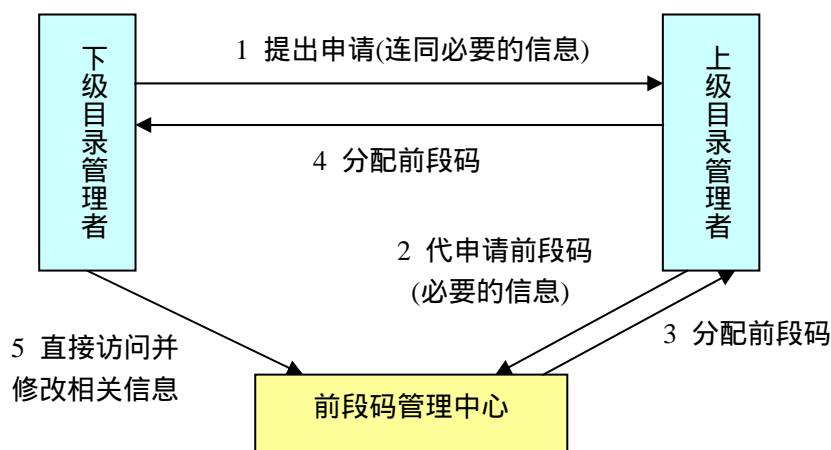


图 2 目录管理者申请前段码的流程

如果没有上级目录管理者，目录管理者可以直接向前段码管理中心提出申请，前段码管理中心认定其符合条件后，可直接为其分配前段码。

政务部门申请前段码时，首先应向本级目录管理者提出，本级目录管理者认定符合条件后，代其向前段码管理中心申请前段码。其申请流程可参见图 2。

如果没有本级目录管理者，政务部门可向上级目录管理者提出申请；如果没有上级目录管理者，也可直接向前段码管理中心提出申请，前段码管理中心认定其符合条件后，可直接为其分配前段码。

政务部门申请前段码，应符合下列条件：

- 正在建立或已经建立了部门的政务信息资源目录体系；
- 具备在本部门范围内管理和分配后段码的能力和设施；
- 其他必要的条件。

由于区县级政务部门一般没有必要建立独立的目录体系，因此原则上不向其分配前段码。

跨市、跨县建立的目录管理者视同地市级或区县级目录管理者。

前段码分配给某一实体后，即使该实体被撤消，也不能将此代码分配给其他的实体。

前段码管理中心参照附录 B 中描述的规则分配前段码。

5.2 后段码的管理与分配

后段码由拥有前段码的目录管理者或政务部门自行管理和分配，必须保证在本部门范围内的唯一性。

后段码原则上采用顺序号，由程序自动生成。编制后段码时应考虑代码的稳定性和简洁性。

拥有前段码的目录管理者或政务部门都可以在编目元数据时生成后段码，以实现政务信息资源的唯一标识。

如果合适，后段码可以采用现有的编码体系。

现有的编码体系必须在一定范围内被权威机构所认可，比如具有稳定的编码规则，并有专人管理。这些编码体系可以是国际或国家范围内认可的编码体系，比如 ISBN 号、ISSN 号等，也可以是一个部门内部为其所管理的信息资源分配的流水号。

5.3 政务信息资源标识符的生成

前段码加上由拥有该前段码的实体为政务信息资源分配的后段码，即构成了一个完整的、全国范围内唯一的政务信息资源标识符。该标识符在政务信息资源的生命周期内保持不变。生成政务信息资源标识符的流程参见附录 A 中的具体案例。

附录 A

(资料性附录)

政务信息资源标识符的生成案例

本附录通过四个具体案例描述了如何生成一个政务信息资源标识符。

案例一：某一目录管理者通过上级目录管理者向前段码管理中心申请前段码

某一信息资源目录管理者 A 在建立目录体系之前，应通过上级目录管理者 B 向前段码管理中心 M 申请前段码，步骤如下：

- 1) A 填写前段码申请表，提交给 B。申请表中应包括机构名称、网址、其他必要信息以及 A 的后段码编码规则。
- 2) B 认定其符合条件后，向 M 申请前段码。M 将按照既定规则生成的前段码传给 B。
- 3) B 将前段码返回给 A，并同时将其保存到自己的系统中。
- 4) A 将前段码加载到自己的政务信息资源标识符生成器（一个程序模块，用来控制并生成信息资源标识符，以下简称 ID 生成器）中，并根据后段码编码规则设置 ID 生成器。
- 5) A 可以登陆 M 的系统，访问自己的相关信息。

某一政务部门通过目录管理者向前段码管理中心申请前段码的步骤同上。

案例二：某一目录管理者直接向前段码管理中心申请前段码

如果没有上级目录管理者，某一目录管理者 A 可以直接向前段码管理中心 M 申请前段码，步骤如下：

- 1) A 填写前段码申请表，提交给 M。表中应包括机构名称、网址、其他必要的信息以及 A 的后段码编码规则。
- 2) M 认定其符合条件后，向其分配前段码。
- 3) A 将前段码加载到本中心的 ID 生成器中，并根据后段码编码规则设置 ID 生成器。
- 4) 如果 A 的上级目录管理者 B 成立，则 M 向 B 告知 A 的相关信息。

案例三：拥有前段码的政务部门向某一目录管理者提交政务信息资源元数据

假定某一政务部门 C 已经申请到了前段码，并准备向某一目录管理者 A 提交政务信息资源的元数据，步骤如下：

- 1) C 对政务信息资源进行机器或手工编目。
- 2) C 利用自己的 ID 生成器生成该资源的标识符。此时，可以分为两种情况：一种是该信息资源原来就具有唯一的内部标识符，编目程序将此内部标识符提交给 ID 生成器，由 ID 生成器自动加上前段码，形成一个完整的政务信息资源标识符，然后向 A 提交。另一种情况是，该信息资源原来没有唯一标识符，ID 生成器按照设定的编码规则自动为其生成后段码，然后加上前段码，形成一个完整的政务信息资源标识符。
- 3) C 向 A 提交元数据。
- 4) A 验证 C 所提交的政务信息资源的标识符元素：首先检查该标识符的前段码是否合法，其次检查后段码是否重复。如果没有异常，则向 C 给出注册成功的响应。否则提示 ID 错误。
- 5) 如果 C 提交的政务信息资源元数据中的标识符元素为空，则 A 利用自己的 ID 生成器为该资源

生成一个标识符，并返回给 C。此时，该标识符的前段码为 A 的前段码，后段码由 A 生成。

案例四：没有前段码的政务部门向某一目录管理者提交政务信息资源元数据

假定某一政务部门 C 没有申请前段码，并准备向某一目录管理者 A 提交政务信息资源的元数据，步骤如下：

- 1) C 对资源进行机器或手工编目，标识符元素为空。
- 2) C 向 A 提交元数据。
- 3) A 利用自己的 ID 生成器为该资源生成一个政务信息资源标识符，并返回给 C。

附 录 B
(资料性附录)
前段码的分配规则

B.1 前段码的分配规则

前段码由信息资源前段码管理中心统一负责分配，原则上按照以下规则进行分配：

1) 前段码的前两位：用于区别不同的省、自治区、直辖市和特别行政区。具体见 B.2。每个省、自治区、直辖市和特别行政区内的所有目录管理者和政务部门前段码的前两位相同。如河北省的前两位为“13”，则河北省范围内的各级目录管理者和政务部门的前段码都以“13”开头。

2) 前段码的第三位：以“0”开头的，分配给省级目录管理者及省级政务部门，比如“130**”分配给河北省目录管理者和省级政务部门。“1”到“Z”开头的前段码分配给省下的各地市，比如“131**”分配给石家庄市。

3) 前段码的第四位：以“0”、“Q”到“Z”开头的，分配给地市级目录管理者和地市级政务部门，比如“1310*”以及从“131Q*”到“131Z*”分配给石家庄市目录管理者和市级政务部门。“1”到“P”开头的前段码分配给地市下的各区县。

4) 前段码的第五位：将“0”分配给区县级目录管理者，“1”到“Z”保留。

前段码的结构和规则具体如图 B.1 所示。

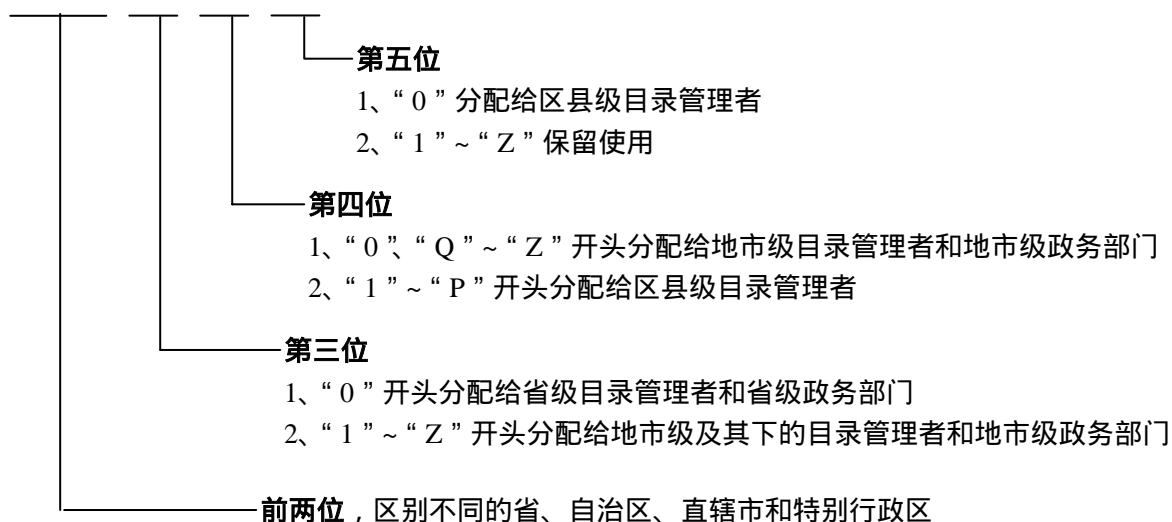


图 B.1 前段码的结构和规则

B.2 前段码前两位的分配

下表给出了前段码的前两位所代表的省、自治区、直辖市和特别行政区。

表 B.1 前段码的前两位

名称	前段码前两位	说明
北京市	11	北京市目录管理者、北京市下属各级目录管理者以及北京市相关政务部门的前段码均以“11”开头。以“11”开头的前段码资源总数为 $36^3=46656$ 。
天津市	12	下同，略。
河北省	13	
山西省	14	
内蒙古自治区	15	
辽宁省	16	
吉林省	17	
黑龙江省	18	
上海市	19	
江苏省	20	
浙江省	21	
安徽省	22	
福建省	23	
江西省	24	
山东省	25	
河南省	26	
湖北省	27	
湖南省	28	
广东省	29	
广西壮族自治区	30	
海南省	31	
重庆市	32	
四川省	33	
贵州省	34	

云南省	35	
西藏自治区	36	
陕西省	37	
甘肃省	38	
青海省	39	
宁夏回族自治区	40	
新疆维吾尔自治区	41	
台湾省	42	
香港特别行政区	43	
澳门特别行政区	44	
