

【GIS 产业化】

美国环境保护署环境及空间数据库系统 USEPA Environmental and Spatial Databases

毛宛红* 李刚†

*AnteonCorp., USEPA Region9, 75 Hawthorne Street
San Francisco, CA 94105

†Department of Geography & Environmental Studies
California State University Hayward, CA 94542

Abstract

In order to better manage large quantity of environmental and spatial information, US EPA has built up Spatial Data Library Systems (ESDLS), Envirofacts and Spatial Data Clearinghouse. ESDLS is a repository for the Agency's new and legacy geospatial data holdings. It enables Agency access to its geospatial data in ARC/INFO format, integrates these data in a standardized, framework. Envirofacts was developed as an enterprise data repository for Agency data systems. It is a relational database implemented in the Oracle Relational Database Management System and contains seven EPA's program system databases. Geospatial Data Clearinghouse provides a pathway to find information about geospatial data used at EPA. Although ESDLS can be accessed through several GIS applications, it is mainly for EPA internal uses. Meanwhile, Envirofacts can be accessed by the public through Internet. Spatial Data Clearinghouse currently provides public access to its metadata. However, it is promised that public can access its geospatial data via the Internet in the future.

美国环境保护署(US Environmental Protection Agency, EPA)是美国政府民用机构中空间(地理)数据的最大用户之一。为了最大限度地利用他们在空间信息资源上的投资,EPA于1987年展开了自己的地理信息系统(Geographic Information Systems, GIS)计划。这个EPA的GIS计划主要由三部分组成,分别是:(1)基础组织机构;(2)空间数据系统;(3)结合GIS和已有的环境数据系统进行环境分析与评价。基础组织机构包括分布于十个EPA大区的GIS中心,位于北卡罗来纳州“研究三角区”内的技术服务,环境监测和项目评价部,以及在弗吉尼亚州总部的一些研究办公室。空间数据系统储存了EPA的环境污染监测,处理及控制设施点的分布,以及一些生态与自然资源的数据。另外,它还结合了美国国家气象站的气象数据,美国人口普查局的普查区,道路及人口数据,和美国地质调查所的地形及其它地理底图。GIS应用分析评价到目前已展开了一些生态系统典型区的环境监测和评估,如在五大湖地区,墨西哥湾及切萨皮克(Chesapeake)湾等进行的将各种时空尺度的环境监测数据转换成生态条件评估和未来自然资源可利用性的预测等。下面将重点介绍EPA的空间数据系统和环境数据系统的数据库结构,数据储存格式,数据内容,数据更新状况,数据提取方式和数据的应用。这两个数据系统是EPA数据管理三大系统中的两个主要组成部分。第三个部分是空间信息交换所。

一. EPA 空间数据系统 (EPA Spatial Data Library Systems, ESDLS)

EPA空间数据系统(ESDLS)是由EPA信息资源管理办公室和EPA地理信息系统计划组负责建立的.ESDLS3.0版是目前最新的EPA国家空间数据系统,它储存了EPA所有的空间数据,可以用来支持各种GIS的应用。

EPA建立ESDLS的基本目标是:建成一个国家范围内统一的空间数据库;将空间数据整体化于一个标准的框架内;应用恰当的标准和指南来管理和使用数据;提供ARC/INFO格式的空间数据给EPA的用户;逐步发展成为环境数据库的一个基本部分。

1. ESDLS 的数据来源

ESDLS的数据来源包括:1990年美国人口普查数据;1992TIGER (Topologically Integrated Geographic Encoding and Referencing);美国地质调查局地理名称信息系统2;水文单元代码;1:2,000,000数字线化图DLG(包括街道,水系,行政界限等);EPA环境数据;环境污染优先治理区界线;主要河流分布;1994年鱼消耗量;邮政编码界限和清单等(请参阅 <http://www.epa.gov/nsdi/pages/title Browse.html>)

在建立过程中，ESDLS 逐渐纳入 EPA 以外的数据。由于这些数据来源繁杂，ESDLS 采用了分步实施的方法。每一步都规划了所有归并数据的地理范围，专题和内容，以及用户服务范围和完成日期。详见表 1。

3). 第三级生态区界线：此生态区界线取自于应用美国大陆的生态区界线和改编的Omernik's framework 生态区界线。这些界线包括：修改过的生态区，定义的亚生态区和一些参考点位。它提供了一个生态空间构架，

表 1。EPA 空间数据系统实施步骤

版本	地理范围	专题内容	用户范围	完成日期
1.1	美国大陆	1992, 1990 年 TIGER 街区, 街区组界线和中心点; EPA 基础设施点; 地理名称点位(GNIS2)数据(医院, 学校, 人类活动区); EPA 超级基金项目(Superfund)区界线; 1:100,000 县, 州界线; 1:2,000,000DLG 道路, 水文, 县州界线	EPA	1996,3,14
1.2	美国大陆	新增加: 主要河流; 1:250,000 水文单元	EPA	1996,3,31
1.3	美国大陆	新增加: Omernik 生态区域; 大气环境信息及提取项目数据	EPA 和其它 联邦机构	1996,8,30
2.0	美国大陆, 阿拉斯加, 夏威 夷, 波多黎各, 以及维尔京群岛	新增加: TIGER92 非美国大陆地区; GIRAS 土地利用; 1994 年鱼消耗情报; 气象站的稳定性数列数据	EPA 和其它 联邦机构	1996,9,30
3.0	美国大陆, 阿拉 斯加, 夏威夷, 波多黎各, 以及 维尔京群岛	新增加: 非美国大陆上的县, 州边界; 1994 年的联邦土地 DLG 数据; 5 位数字的邮政编码边界(美国大陆, 阿拉斯加, 夏威夷)	EPA 和其它 联邦机构	1997,9,30

2. ESDLS 的数据

空间数据系统的空间数据采用 ARC/INFO 格式，有标准的和统一的数据使用和管理规则。这些空间数据按地区等级分为国家级，州级和县级。比例尺分别是：1: 10 万(县级)，1: 25 万(州级)，1: 200 万(国家级)。地理区域包括：美国大陆，阿拉斯加，夏威夷，波多黎各，以及维尔京群岛。数据涉及的主要专题内容包括：

1). EPA 常规设施点位分布：反映了由 EPA 掌握的主要环境监测，处理，控制设施点的分布。这些点位数据主要取自于 EPA 的环境媒体及其它研究项目，详见本章第二部分。这些数据可用于多种分析，例如：人口和设施的比邻关系等。目前，这些数据用于 EPA 生态方面的研究和万维网上的各种 GIS 应用。

2). 主要河流分布：它包括了美国大陆上将近 700,000 公里的水系。这个数据来源于美国地质调查局 1: 250,000 地形图上的所有水文要素，并在其基础上加工更正充实。它完成于 1982 年，已经多年被用于 EPA 和国家的鱼类，野生动物以及气象服务。

可用于环境资源的研究。

4). 1994 年鱼类受污染信息：它包括了化学污染情况，受污水域，受污染鱼种和数量，以及污染的开始与结止日期，持续时间等。它向公众揭示了一系列水域和鱼种的潜在污染危机。

5). 美国气象站的气象数据：它总结了国家气象站的气象数据，其中包括：六个风速的综合频率，十六个风向，以及六个稳定性等级。这些要素取自于各气象站年报中每小时的记录。本数据用于 EPA 的图形研究模型系统。

6). 1992 年美国人口普查局的 TIGER 空间数据：包括道路，铁路，水文，点，线，面等。这些线划数据是根据 1990 年人口普查的表格数据和公开发表的边界由县或同等区级发布的。EPA 的国家地理信息服务中心就这些数据处理成 ARC/INFO 格式。此类数据可用于作图，路线分析，人文地理分析，以及 EPA 设施布局与人口分布的关系分析。

7). 1995 年美国地质调查所地理名称点位：此数据所

含有的地理名称是以实际地物的经纬坐标来标识的。本数据通常用作地理底图。

8). 1990 年美国人口普查局的街区界限, 街区中心点, 街区组界限等: 这些数据取自于 1990 年美国人口普查局的 TIGER 线划数据。此线划数据提供了 1990 年人口普查所用的普查区边界。

9). 1:2,000,000 比例尺的州际公路, 水文, 州界及县界: 这些数据取自于 1991 年美国地质调查局 1: 200 万数字线划图 DLG 数据。

10). 美国超级基金项目点位: 反映了 1,223 个国家优先考虑的综合环境污染点的分布区。

11). 美国和墨西哥边界 100 公里范围内的大气污染监测点: 美国和墨西哥边界 100 公里范围内从 1983 年起被列入环境关注区。本数据被美国—墨西哥边界项目所用。

12). 1:250,000 GIRAS (Geographic Information Retrievaland Analysis System) 土地利用和覆盖信息: 此数据是由美国地质调查所收集, 由 EPA 转换为 ARC/INFO 数据。由于数据的采集经历了从 70 年代到 80 年代 10 多年的时间, 数据详细程度差异比较大, 土地利用分类采用 Anderson 的分类方法。本数据所反映的土地利用组合形态曾应用于与水质有关的环境评价, 农业生产管理, 以及其它环境影响评价分析。

此外还有各种关于大气, 水文, 土地环境状况的信息, 这里不一一介绍。除了现有的数据以外, 将来准备增加新的专题数据, 其中包括: 水文单元类型索引系统; 国家公园边界; 国家野生动物保护区; 濒临灭绝物种分布等等。

二. 环境数据系统 (EnviroFacts)

环境数据系统储存了 EPA 的环境设施, 项目及观测结果数据。环境数据系统储存的数据取自如下 EPA 的七个信息系统和数据库:

1). 综合环境治理, 补偿, 责任追究信息系统: 这个项目是针对美国多年来积累的大量废弃物无人处理的问题由参议院于 1980 年设立的, 是 EPA 超级基金 (Superfund) 的一部分。此项目的目的是彻底清除大型灾害性废弃物。这个项目积累了从 1983 年至今的大型灾害性废弃物评价, 清除计划, 基金投入, 和所用技术等等信息。

2). 排污许可检控项目: 美国所有的污水排放, 不论是城市污水还是工业污水排放, 都需要事先获得排放许可。排污许可检控项目是根据环境数据来检查和调整排放许可的上线以及排放许可的多少。

3). 资源保护信息系统: 这个系统提供了灾害性废弃物信息, 用来支持资源保护的措施。这个项目要求灾害性废弃物的制造者, 运输者, 储存者, 处理者向州级环保部门提供他们的活动数据。州环保部门再将信息汇总到 EPA 大区和总部。所提供的数据内容有: 废弃物的物主资料; 废弃物的处理方法, 废弃物处理的监视程度, 废弃物处理的跟踪调查等等。

4). 有毒物质排放清单系统: 这个系统所提供的数据包含了由制造厂每年向空气, 水, 土地释放的大约 300 种有毒物质。工业设施所提供的信息包括: 工厂生产, 处理, 使用化学品的设施的位置; 化学品的储备量; 化学品的大概排放量; 化学品就地再加工和再循环的能力; 化学品运往它处进行处理, 再加工, 再处理的大概数量等。

5). 空气信息提取系统及大气点污染源信息系统: 这个系统记录了几乎所有的美国点源污染点的资料, 包括工厂类型, 位置, 污染物排放点和排放量, 以及工厂内部生产过程和可能造成的污染。这些资料由空气信息提取系统和大气污染设施子系统(Aerometric Information RetrievalSystem/AIRS Facility Subsystem, AIRS/AFS) 来储存, 管理和提取。

6). 安全饮用水信息系统: 这个系统所储存的安全饮用水信息包括公用水系统和违反 EPA 安全饮用水规定的记录。EPA 以法律形式规定了安全饮用水的标准, 水处理技术, 水质监测等等以确保人民的健康。

7). 研究与治理基金的管理信息系统: 这个系统存有所有 EPA 项目的基金申请, 批准和管理信息。

此外, 环境数据系统还支持三个横向关系数据库:(1) EPA 设施检索系统,(2) 主要化学物质综合数据库, 以及(3) 点位数据库。其中 EPA 设施检索系统从横向联系了多个数据库中的 EPA 设施。主要化学物质综合数据库为排污许可检控系统, 综合环境响应及补偿系统和有毒物质排放清单系统提供了化学物质横向联系的标准。点位数据库以经纬度坐标形式描述了 EPA 设施的地理位置。

1. 数据库构成及数据储存

环境数据系统是应用 ORACLE 关系数据库建立的。它将上述 EPA 的七个各自独立的数据库和三个横向关联的数据库联在一起。只要进入这个系统就可以提取到上述十个数据库的任何信息。图 1 表示了各数据库间的关系。

2. 数据查询

EPA 已经在万维网上建立了多种环境事实数据查询和制图功能。用户可以有选择地查询某个 EPA 项目数据库, 也可以同时查询所有的 EPA 项目数据。查询方式有: 按

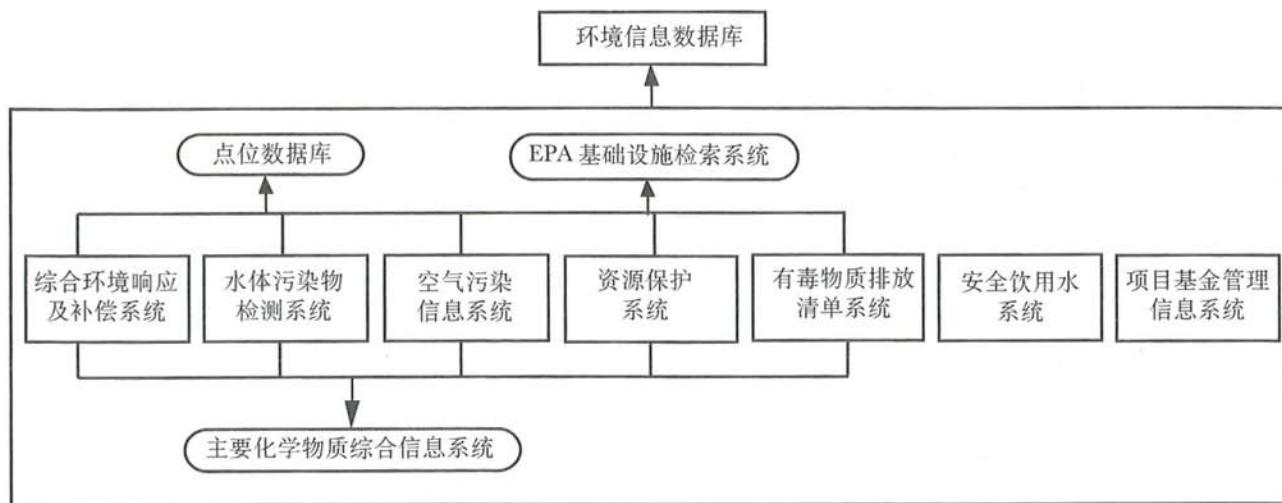


图 1. 环境信息数据系统数据结构

EPA 设施名称或代号, 地理区域, 标准工业分类代码, 或按照数据库中的某个数据元素。数据查询的结果是一个列有符合用户要求的 EPA 设施清单。

3. 数据更新

环境数据系统基本上是按月更新。其中, 研究与治理基金的管理信息系统是每两周更新一次。安全饮用水信息系统是每三个月更新一次。对于按月更新的 EPA 信息系统, 今后环境数据系统将增加新的 EPA 信息子系统, 其中包括: 紧急事件通报系统和综合环境响应及补偿系统。

这个数据系统对 EPA 内部用户和公众免费开放, 其目的是一方面使得公众可以接触到各种 EPA 的数据, 另一方面是为 EPA 用户提供了万维网和其它 ORACLE 工具用以检索数据。

三. EPA 地理空间数据交换所(Spatial Data Clearinghouse)

国家地理空间数据交换所 EPA 交换点是国家空间数据基础构架的一部分。EPA 交换点的运行是在技术服务中心(Enterprise Technology Services Division, ETSD)的支持下通过信息资源管理办公室(Office of Information Resources Management, OIRM)的国家地理信息系统来完成的。这个交换点为公众开辟了一条通往 EPA 空间数据的渠道。

国家地理空间数据交换所的目的是让用户了解现有的数据, 找到所需的数据, 评价数据的可应用性, 以及定购数据。EPA 交换点存有 50 多种元数据, 面向公众开放。它的万维网地址是: <http://www.epa.gov/nsdi>。

EPA 的元数据是通过 EPA 总部及其分部的 GIS 活动获

取来的。这些元数据需要参照联邦地理数据委员会(Federal Geographic Data Committee, FGDC)标准进行检查和补充。由于这个查证过程比较复杂,EPA 的公关技术服务部(Enterprise Technical Services Division)专门在北卡罗来纳的“研究三角园”设立了一个中心以从事元数据的检索, 上载及传播。

EPA 交换点根据需要每个月增加一些新的元数据。它们大多数按照 EPA 设施的编号进行检索, 查询, 和提取。这个交换点同时保存了没有进行检索处理的(non-indexed)的 EPA 空间数据源和其它地理空间数据源, 并且建立了连接联邦, 各州, 大学, 其它国家, 商业, 和其它国家空间数据基础设施(National Spatial Data Infrastructure, NSDI)信息站的通道。

四. 空间信息系统与环境信息系统的结合—按需制图(Map on Demand)

按需制图功能应用于将环境数据库的数据按各种用户需求作图。这个制图系统采用了所有的 EPA 数据, 其中包括 EPA 环境信息系统的数据, EPA 空间信息系统的数据和人口统计数据。因而用户可以有多种多样的制图选择。目前, 按需制图的功能主要有: 定点制图, 流域制图, 县域制图, 邮电编码制图, EPA 设施密度制图, 以及资料查询结果制图。

1). 定点制图功能是将用户按经纬度选择的地点上的人口统计信息显示出来。制图区域是由用户设定的半径大小所决定。用户可以按照环境数据系统中的查询表来选择经纬度。这个功能最初被 EPA 的十个大区分部用来帮助环境决策。表 2 和图 2 是这个功能的实例。

2). 流域制图功能是按用户要求显示人口统计和 EPA 信息系统的数据在流域中的分布。同时还以文字报告的形式给出 EPA 常规设施在水文单元中的信息。

表 2. 按需制图一定点制图申请表

地图主标题	<input type="text"/>		
地图副标题	<input type="text"/>		
制图区域	度	分	秒
纬度	<input type="text"/>	<input type="text"/>	<input type="text"/>
经度	<input type="text"/>	<input type="text"/>	<input type="text"/>
制图区域半径或比例尺			
图幅 14x11 平方英寸, 制图比例尺取决于所选择的半径大小。 可选择的半径有:			
<input type="checkbox"/> 2 公里 <input type="checkbox"/> 5 公里 <input type="checkbox"/> 10 公里 <input checked="" type="checkbox"/> 10 公里和 2 公里 (两张图) <input type="checkbox"/> 5 公里和 2 公里 (两张图)			
图幅 24x24 平方英寸, 制图比例尺为 1:24000			
选择所需的文字报告			
EPA 常规基础设施信息			
<input type="checkbox"/> 水中污染物排放量的监测设施及数据 <input type="checkbox"/> 有毒物质排放的监测设施及数据 <input type="checkbox"/> 综合环境响应及补偿系统设施 <input type="checkbox"/> 资源保护系统设施			
人口、种族统计信息			
<input type="checkbox"/> 公用水供给信息 <input type="checkbox"/> 敏感环境条件信息			
选择背景图信息			
<input type="checkbox"/> 人口密度 <input type="checkbox"/> 有色人口百分比 <input type="checkbox"/> 贫困人口百分比			
选择一个或多个专题信息			
<input type="checkbox"/> 公路和铁路 <input type="checkbox"/> 水网 <input type="checkbox"/> 湿地 <input type="checkbox"/> 国家公园及消闲地 <input type="checkbox"/> EPA 授权的水流量检测设施 <input type="checkbox"/> EPA 超级基金项目设施 <input type="checkbox"/> 资源保护及恢复项目的设施 <input type="checkbox"/> EPA 有毒物质释放的检测设施 <input type="checkbox"/> 公用地下水供给 <input type="checkbox"/> 经纬网			
地图输出格式			
用于万维网看的图形数据格式:			
<input type="checkbox"/> GIF 低分辨率 (72dpi) <input type="checkbox"/> GIF 中分辨率 (144dpi) <input type="checkbox"/> GIF 高分辨率 (261dpi) <input type="checkbox"/> AcrobatReaderPDFPlotfile			
用于非万维网的地图及打印文件格式:			
<input type="checkbox"/> Arc/InfoArcplot.gra.zipfile <input type="checkbox"/> GeneralColorPostscriptPlotFile.ps.zip <input type="checkbox"/> AdobeIllustratorPlotFile.ai.zip <input type="checkbox"/> HPGL2-8PenDraftmasterPlotFile.hpgl.zip <input type="checkbox"/> WindowsBitmapFiles.bit.zip <input type="checkbox"/> ...			
你的 e-mail 地址: <input type="text"/>			
<input type="button" value="发出作图要求"/> 或 <input type="button" value="重新选择参数"/>			

3). 县域制图功能是制作以县为单位的有关EPA管理和健康问题的图件及报告。例如, EPA常规设施, 人口统计, 和安全饮用水在用户指定县中的分布。
4). 邮电编码制图功能和县域制图功能相似。所不同的是, 前者是以邮电编码为制图单元。

5). EPA设施密度制图是编制以国家, 州, 县, 及水文流域为单元的EPA常规设施图。用户可以从这类图上了解到EPA常规设施的集中分布, 同时, 也从所提供的文字报告中掌握这些设施的数量情况。

6). 查询结果制图是将环境事实数据系统查询结果制图。图上除了环境信息以外还显示了制图区域的人口统计信息, 土地利用, 土地覆盖, 自然和文化特征。图上面状符合可表现为土地利用状况, 人口密度, 少数民族百分比, 贫困人口百分比, 人均收入。查询结果制图容许多层次地理数据制图。图型文件的格式有: GIF, TIFF, JPEG等的压缩格式。

五. 小结

本文对EPA环境与空间数据系统进行了概述。目前这些系统仍不够完善, EPA对这些系统及它们内部的各个部分仍在进行改进。比如, 空间数据系统(ESDLS)和地理空间数据交换所(SpatialDataClearinghouse)还不能对公众完全开放。空间数据系统现只对EPA内部用户开放。公众可以查询地理空间数据交换所的元数据(Metadata), 但还不能通过它直接调用空间数据。另外, 为了提高设施点位数据的精度, EPA从1990年开始进行点位数据更正项目。这个项目的目的是为了确保数据的精度, 格式的一致性和数据的完整性。它规定将精度保证在(25米。因此, 采集EPA设施位置的精确经纬度是这个项目的目标。数据采集的对象主要包括: 所有EPA常规设施, EPA项目活动点和环境监测观测点。除了经纬度坐标以外, 其它需要收集的相关数据有: 经纬度坐标的测量方法, 测量精度, 地物实体, 元数据比例尺等等。为了更加规范地记录所需数据, EPA特别制定了“方法, 精度和描述”参照表来指导数据的正确表达。

参考文献

- US Environmental Protection Agency, Envirofacts Database Description Document Version 2.0, SDC-0055-051-MA-5031, September 30, 1996.
- US Environmental Protection Agency, ESDLS Version 2.0 Operations and Maintenance Manual, SDC-0055-065-DY-5039B, August 22, 1996.
- US Environmental Protection Agency, Design Document for the Geospatial Data Component of ESDLS Version 2.0, SDC-0055-065-KF-5034, September 30, 1996.
- US Environmental Protection Agency, Design Document for Locational Reference Tables in Envirofacts Version 2.1, SDC-0055-091-RP-6013, January 21, 1997.

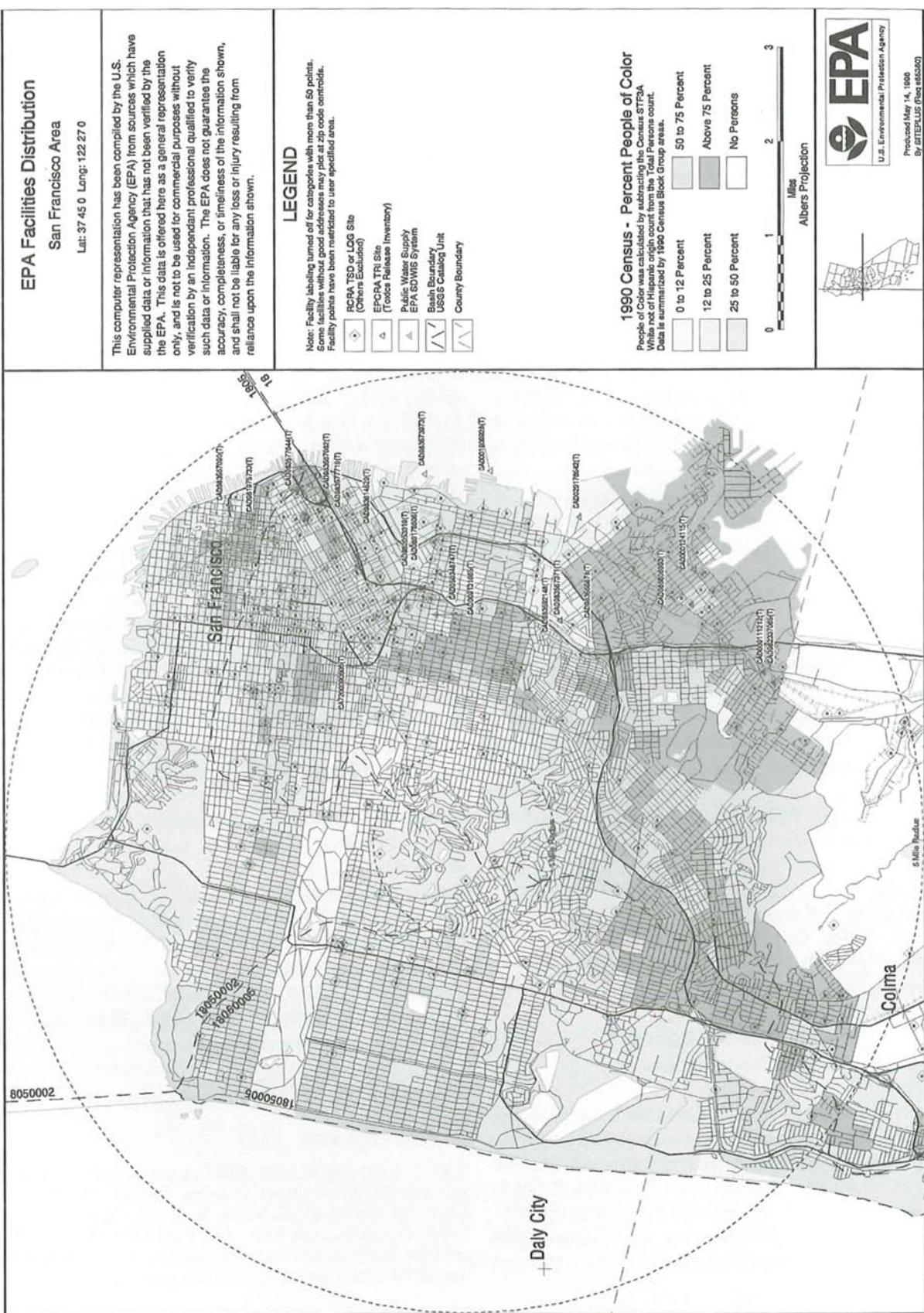


图 2. 旧金山市 EPA 设施分布图—按需制图功能的一个实例