

中国毒品、犯罪和拘留数据库项目

项目概览

阿米尔·德克斯

2019年5月28日

数据库简介：

中国毒品、犯罪和拘留数据库项目是第一个专注于中国政府毒品政策和国家司法系统的开放数值型数据库。该数据库使用了200多个公开数据来源，涵盖500多个子类别，涉及议题包括：非法药物使用、减少危害干预措施、刑事和行政犯罪、司法审判和治安等。

作为开放数据资源，该项目解决了中国问题研究者面临的共同困境：如何获取数据。相较以往，中国现在公开了更多关于毒品、犯罪和司法审判的数据。但这些官方数据往往分散在难以计数的信息来源中，碎片化且不稳定。例如，一些早年国家政府报告中包含的数据类别在近年报告中消失了，

有些数据仅出现在新闻报道或学术著作中。虽然这些数据是公开的，但从中收集、整理信息需要大量的时间和精力。通过集中梳理这些分散的数据，该数据库为解决这些问题。

该数据库由伦敦政治经济学院(LSE)国际药物政策研究中心研究助理兼多伦多大学政治科学博士生阿米尔·德克斯创建。伦敦政治经济学院发展研究项目研究生徐一斐女士对该项目提供了研究和翻译支持。该项目在发布前通过了伦敦政治经济学院伦理委员会的伦理审查，并通过DMPonline建立了数据管理计划。

数据库使用：

作为开放数据资源，该数据库供研究中国毒品政策、刑事司法、治安和拘留问题的研究人员、倡议人士、记者和相关组织使用。用户可以通过数据库检索一系列议题，包括：

- 中国境内六种以上刑事和法外拘留场所的拘留总人数；
- 刑事和行政犯罪的逮捕率；
- 司法审判模式，涵盖危害国家安全罪行的近期数据；
- 禁毒系统对社区矫正、缓刑和假释措施的加强使用趋势；
- 中国毒品市场从海洛因主导到甲基苯丙胺主导的变化趋势；

在保留戒毒拘留措施的同时，加强使用减少危害干预措施；
逐渐上升的登记在册毒品使用者的人口统计数据；
公安部和人民武装警察部预算比例；
中国与美国等其它国家的数据对比。

目前数据库仅呈现了国家层面的数据，未来将为中国23个省、5个自治区和4个直辖市建立单独的子数据库。这些子数据库将有助于研究人员分析一系列数据指标在中国境内的地域差异。

数据来源：

数据库数据均来自可公开访问的中英文资源，包括：

中国公安部年度禁毒报告；
中国国家统计局年鉴；
中国法律年鉴；
媒体报道；
学术出版物；
倡议团体、外国政府、国际组织和国际非政府组织的报告。

数字化信息通过百度、谷歌、中国知网和JSTOR获取。非数字化信息通过多伦多大学图书馆系统获取，或从研究员个人收藏中获取。

数据类别：

数据库涵盖了1990年至今中国与非法药物、犯罪、司法审判、拘留和公共安全有关的数据，亦涵盖了相关人口统计和经济数据。这些类别的选取首先基于对公安部、国家统计局发布的报告和中国法律年鉴中相似数据类别的统计而完成。数据库中的其它数据类别均由研究员创建。将来可能为响应用户兴趣而创建新数据类别。

补充文件：

用户可以通过Zenodo下载数据库的补充文件。文件均为中英文对照，包括：

数据库500多个数据类别清单；
存在差异的数据类别的不同预估结果和相关解释；

参考数据来源。

数据处理：

鉴于数据库使用了大量信息来源，数据差异不可避免。上述补充文件提供了存在不一致的数据类别的不同估算结果。用户可以使用该文件中的替代数据替换数据库中的原有数据。数据库中的空白代表未能获取的数据，但该数据可能存在。

数据开放承诺：

作为开放数据资源，该数据库和相关文件可以通过Zenodo和LSE国际药物政策研究中心网站免费下载。使用数据时，用户只需正确引用数据库即可。

该数据库的开放承诺涵盖用户参与数据库的改进。如果用户发现数据错误、能提供新数据或有任何疑问，请联系研究员阿米尔·德克斯：电子邮件：emile.dirks@mail.utoronto.ca；

Twitter：@emiledirks