

全球大学纵向招生数据库

伊丽莎白·巴克纳

伊丽莎白·巴克纳 (Elizabeth Buckner)：加拿大多伦多大学 (University of Toronto) 安大略教育研究院 (Ontario Institute for Studies in Education) 助理教授

电子邮件: elizabeth.buckner@utoronto.ca

全 20 世纪 70 年代，全球私立大学（属于国际教育标准分类第六级，即学士或等同水平）的规模有多大？最近私立高等教育的增长主要是因为新私立大学的成立，还是已成立院校机构的扩招？这些只是私立高等教育学者为更好地理解全球私立高等教育的发展及其对高等教育格局、政策和学生的影响而提出的一些问题。

然而，直到现在，高等教育领域的学者很难将近年来的增长与早期时代进行对比。这是因为缺乏 2000 年以前私立高等教育的数据。本文讨论了大学招生数据的许多问题，并介入了一个新的数据集，旨在推动对大学的跨国、跨时期研究。

缺乏按部门划分的高等教育招生数据

现有的私立高等教育跨国研究数据存在诸多局限性。联合国教科文组织统计研究所 (UNESCO's Institute for Statistics, 简称 UIS) 从 1998 年开始收集全球私立高等教育入学率的数据。然而，1998 年只有 14 个国家的数据。截至 2005 年，有 72 个国家的私立高等教育入学率数据。UIS 的数据覆盖范围在 2015 年达到峰值，当时有 127 个国家的估计数，之后数据缺口再次增大，很可能是由于数据报告的自然滞后。到 2020 年，只有 54 个国家的估计数据可用。UIS 数据的另一个问题是，它依赖于国家或联邦政府的报告，

这意味着有些国家的数据永远无法获得，包括我自己所在的国家加拿大。

此外，教科文组织对“公立”和“私立”的定义并不总是符合国家的定义或公众对公立和私立高等教育的理解。这对英国等国家的后果是显而易见的，在英国，私立大学的入学人数在所有年份都占有所有类型大学入学人数的 100%。

因此，对提升私立高等教育统计数据准确性感兴趣的学者们转向了“私立高等教育项目” (Program on Private Higher Education) 等数据来源，这些数据更符合大众的理解和政府政策。“私立高等教育项目”的网站提供了基于 UIS 的详细国家统计数据的估计，并在无法获得政府数据的情况下提供政府数据来源和国家级专家。然而，“私立高等教育项目”的一个限制是，到目前为止，只有 2010 年的数据是公开的，这使得跟踪长期趋势变得困难。

现有数据集的第三个局限性是，关于私立高等教育的现有数据将短周期高等教育（国际教育标准分类第五级）与大学（国际教育标准分类第六级及以上）结合在一起，即提供学士和/或高级学位的机构。然而，高等教育数十年的学术研究已经证明了大学的独特特征，它们更有可能以科学研究为导向，参与全球排名，而且往往享有更多的学术自由。对于我们这些对大学作为一种社会机构

特别感兴趣的人来说，短周期高等教育机构与大学的合并是一个主要问题。

与私立高等教育相关的问题得到了更广泛的高等教育研究人员的认同：大多数公开的高等教育数据都是在国家层面上计算的，将短周期招生和大学招生结合在一起，而且直到最近几年才被广泛使用。

介绍 GLUED 的测试版

显然，我们需要更好地了解大学各部门长期的招生情况。全球大学纵向招生数据集（The Global Longitudinal University Enrollment Dataset，简称 GLUED）是一个新的数据集，旨在做到这一点。GLUED 的三个主要优势是：它代表了全球大学（国际教育标准分类第六级及以上）的近普查；这些数据是在院校层面收集的，这将使研究人员能够调查院校组织层面的现象；对许多院校来说，从 1950 年开始，原始和估计的招生数据都是可用的。

构建 GLUED 花了很多年。这些数据由多伦多大学安大略教育研究院收集，并得到了加拿大社会科学与人文学研究理事会（Social Science and Humanities Research Council）的慷慨资助。在 2018 年至 2020 年期间，我们的研究助理团队手动输入了数千所大学的招生数据，这些数据来自欧罗巴学习世界（the Europa World of Learning）和国际大学协会（International Association of Universities）的世界高等教育数据库。这些数据补充了来自美国综合高等教育数据系统（Integrated Postsecondary Education Data System）的数据，来自国际大学协会主办的世界高等教育数据库（数字版本）的最新入

学数据，以及来自维基百科的学生入学数据。所有数据都经过了全面检查和清理。根据特定院校的增长率或院校特征估计了前几年的缺失数据，并对其进行校准，以与全球和国家两级公布的数据来源保持一致。

GLUED 列出了 185 个国家和地区 15263 所院校的校名和特点，包括每隔五年估计的学生入学数据。另一个国家层面的数据，包括到国家一级的具体部门的入学数据，同样能够在 GLUED 中获取。

GLUED 还包括一些从各种来源合并的关于院校特征的其他有用变量，包括部门（即公立/私立）性质、成立年份，以及机构是否授予博士学位。大学部门性质（公立/私立）是每所学校根据机构控制或所有权的原始来源自行定义并宣布的，而 GLUED 采用了这一定义。GLUED 数据集没有区分营利性或非营利性大学，也没有捕捉到一所大学是否有宗教信仰，但这是未来版本的数据集可能会添加的内容。

最后，GLUED 的一个令人兴奋的功能是，它还包括从谷歌地图上抓取的每所大学的地理位置（即 GPS 坐标）。这些数据可用于在数据集中绘制大学的地理位置图，并使世界各地大学的发展历程可视化，同时还可被科研人员用于广泛目的的研究工作。

该数据集目前还处于测试版本的阶段，我们也在在此期间完成了数据集清理和缺失值估计。该数据集将于 2022 年秋初在多伦多大学的数据平台上发布，并将通过加拿大数据平台 Borealis 访问，访问地址如下：<https://doi.org/10.5683/SP3/P0D1KE>。